







Statistica piețelor financiare

Seminarul 4: Probabilitate



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

- **AI**: Artificial Intelligence (inteligență artificială)
- **AR**: AutoRegressive (model) — model autoregresiv
- **BET**: Bucharest Exchange Trading (index) — indicele de referință al Bursei de Valori București
- **CDF**: Cumulative Distribution Function (funcția de repartiție)
- **EODHD**: EOD Historical Data (market-data provider) — furnizor de date de piață
- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity (heteroscedasticitate condiționată autoregresivă generalizată)
- **GBM**: Geometric Brownian Motion (mișcarea browniană geometrică)
- **LCG**: Linear Congruential Generator (generator congruențial liniar)
- **PCG64**: Permuted Congruential Generator, 64-bit (NumPy default) — generatorul congruențial permutat pe 64 de biți (implicit în NumPy)
- **PMF**: Probability Mass Function (funcția de masă de probabilitate)
- **RANDU**: RANDU (IBM random number generator of the 1960s) — generatorul de numere aleatoare RANDU al IBM, anii 1960

Întrebarea de azi și structura seminarului

- Întrebarea seminarului
  - **cum descriem, condiționăm și simulăm randamente și prețuri aleatoare?**
- Exercițiile: în ordinea din Harta exercițiilor (slide-ul următor)
- Rezultate de obținut
  - răspunsurile la exercițiile A; tabelul sau figura cerute la fiecare exercițiu B
  - un răspuns scurt și argumentat la C1; erorile găsite în răspunsul AI din C2 și răspunsul corectat
  - frazele de interpretare cerute la fiecare exercițiu
- Seminarul are rol de exercițiu și nu se notează; rezolvările exercițiilor [Propus] se discută la seminar
- Notebook-ul de azi: deschideți notebook-ul seminarului în Google Colab; fiecare cerință indică secțiunea
- Contact: prof. Daniel Traian Pele, danpele@ase.ro

Harta exercițiilor

Cerința	Întrebarea	Tipul	Model
A1, A2	momentele unui randament discret; este independentă o pereche necorelată?	Rezolvat, Propus	A1
A3, A4	volatilitatea unui portofoliu; un model cu două regimuri și legea variației totale	Rezolvat, Propus	A3
A5, A6	un arbore binomial; mers aleator, martingal și AR(1)	Rezolvat, Propus	A5
A7, A8	transformarea inversă pas cu pas; GBM și procesul Wiener	Rezolvat, Propus	A7
B1, B2	sînt randamentele independente? media și varianța condiționate	Rezolvat, Propus	B1
B3, B4	probabilitatea unei pierderi prin Monte Carlo; testarea unui generator de numere aleatoare	Rezolvat, Propus	B3
B5, B6	GBM față de BET; crește varianța liniar cu orizontul?	Rezolvat, Propus	B5
C1, C2	cît de des apare un an nefavorabil? ce este greșit într-un răspuns AI?	Propus	B3, B5

- **[Rezolvat]**: rezolvarea completă în slide-uri și în notebook, un model de urmat; **[Propus]**: îl rezolvați dumneavoastră, după model
- **A**: calcule pe hîrtie; **B**: date reale, în notebook; **C**: întrebări deschise și critica unui răspuns generat de AI

Datele folosite

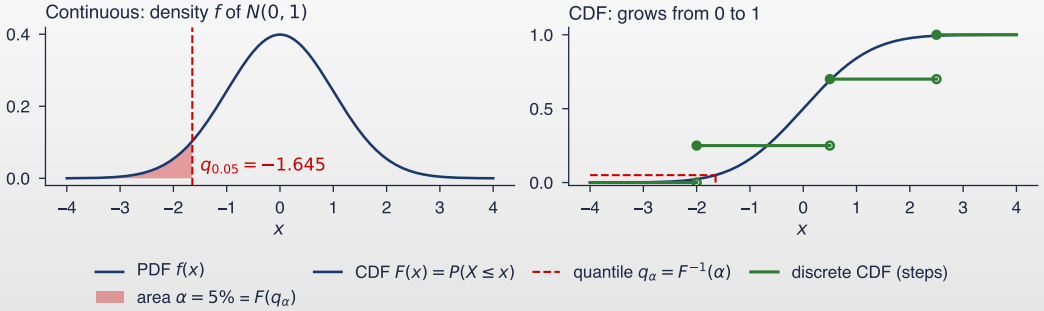
Seria	Sursa	Prețul folosit	Perioada
S&P 500, DAX, BET	EODHD	close	2000–2026
Bitcoin	EODHD	close, 7 zile pe săptămînă	2014–2026

- Randamente logaritmice zilnice în %: $r_t = 100\,(\ln P_t - \ln P_{t-1})$
- Weekendurile și închiderile repetate din zilele libere se elimină (cu excepția Bitcoin); fiecare serie își păstrează propriul calendar
- Simulările folosesc generatorul implicit din NumPy cu semințe fixe, astfel încît fiecare rezultat numeric poate fi reprodus

Noțiuni necesare azi (1/21): distribuții

- O **variabilă aleatoare** X: un număr determinat de hazard (randamentul de mîine)
  - X discretă: **PMF** (probability mass function, funcția de masă de probabilitate) $p(x_i) = P(X = x_i)$, probabilitatea fiecărei valori x_i
  - X continuă: **PDF** (probability density function, densitatea de probabilitate) f; probabilitatea unui interval este aria de sub f
- **CDF** (cumulative distribution function, funcția de repartiție): $F(x) = P(X \leq x)$
  - probabilitatea ca X să nu depășească x; crește de la 0 la 1
- **Cuantila** de ordin α: $q_\alpha = F^{-1}(\alpha) = \inf\{x : F(x) \geq \alpha\}$
  - cel mai mic x cu $F(x) \geq \alpha$; mediana este $q_{0,5}$

Noțiuni necesare azi (2/21): densitatea, funcția de repartiție și cuantila pe un grafic



- Stînga: aria de sub PDF între a și b este probabilitatea acelui interval

$$P(a < X \leq b) = \int_a^b f(x)\,dx = F(b) - F(a)$$

- Dreapta: o funcție de repartiție continuă este netedă; una discretă este în trepte, cu saltul $p(x_i)$ în x_i
  - cuantila q_α: de la α orizontal pînă la curbă, apoi în jos; pentru $N(0,1)$, $q_{0,05} = -1{,}645$

SFM_ch4_seminar

Noțiuni necesare azi (3/21): media și varianța

- **Speranța matematică**: $E[X] = \sum_i x_i p(x_i)$ (discret) sau $\int x f(x)\, dx$ (continuu)
  - media valorilor, ponderată cu probabilitățile lor
- **Varianța**: $\mathrm{Var}(X) = E[X^2] - (E[X])^2$
  - distanța pătratică medie față de medie; abaterea standard $\sigma = \sqrt{\mathrm{Var}(X)}$, volatilitatea unui randament
- **Boltirea**: $E[(X-\mu)^4]/\sigma^4$, cu $\mu = E[X]$
  - este 3 pentru orice distribuție Normală; peste 3: cozi mai groase (excesul de boltire = boltirea -3)
  - pentru $N(0, \sigma^2)$: $E[X^4] = 3\sigma^4$

Noțiuni necesare azi (4/21): momentele de selecție

- Din randamentele observate $r_1, \ldots, r_n$ estimăm momentele

$$\bar{r} = \frac{1}{n}\sum_{t=1}^{n} r_t, \qquad \hat{\sigma}^2 = \frac{1}{n-1}\sum_{t=1}^{n}(r_t - \bar{r})^2$$

  - $\bar{r}$: media de selecție, o estimare a lui μ; $\hat{\sigma}$: abaterea standard de selecție; notația cu „căciulă" indică o estimare

- **Boltirea de selecție** și excesul de boltire

$$\hat{K} = \frac{\frac{1}{n}\sum_t (r_t - \bar{r})^4}{\left(\frac{1}{n}\sum_t (r_t - \bar{r})^2\right)^2}, \qquad \text{excesul de boltire} = \hat{K} - 3$$

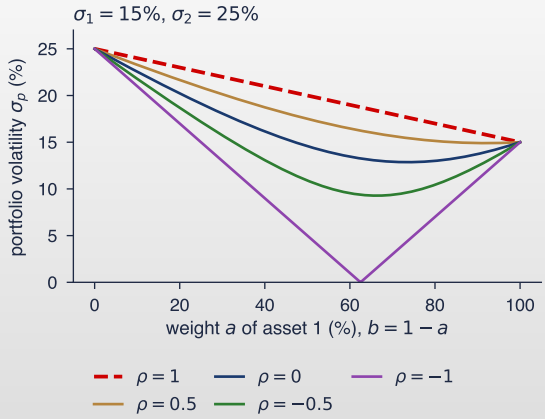
  - aproape de 0 pentru date Normale; un exces mare și pozitiv: mai multe randamente extreme decît dă distribuția Normală

- $\mathrm{Corr}(r_t^2, r_{t-1}^2)$: corelația de selecție a perechilor (r_t^2, r_{t-1}^2), $t = 2, \ldots, n$
  - aproape de 0 pentru randamente i.i.d. (independente și identic distribuite); pozitivă cînd mișcările mari urmează după mișcări mari (volatility clustering)

Noțiuni necesare azi (5/21): covarianța și portofoliile

- **Covarianța**: $\mathrm{Cov}(X, Y) = E[XY] - E[X]E[Y]$
  - pozitivă cînd X și Y tind să se miște în același sens
- **Corelația**: $\rho = \mathrm{Corr}(X, Y) = \mathrm{Cov}(X, Y)/(\sigma_X \sigma_Y) \in [-1, 1]$
  - σ_X, σ_Y: abaterile standard; ρ măsoară doar dependența liniară
- **Varianța unei sume**: $\mathrm{Var}(aX + bY) = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\,\mathrm{Cov}(X, Y)$
  - a, b: ponderile din portofoliu; termenul de covarianță produce efectul de diversificare cînd $\rho < 1$

Noțiuni necesare azi (6/21): diversificarea pe un grafic



- Portofoliu cu ponderile a și $b = 1 - a$

$$\sigma_p = \sqrt{a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2}$$

  - σ_1, σ_2: volatilitățile celor două active; ρ: corelația lor
- $\rho = 1$ (linia întreruptă): $\sigma_p = a\sigma_1 + b\sigma_2$, media ponderată
  - orice $\rho < 1$ dă o curbă sub această dreaptă: cu cît ρ este mai mic, cu atît cîștigul este mai mare
- Ponderea de variantă minimă

$$a^* = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$$

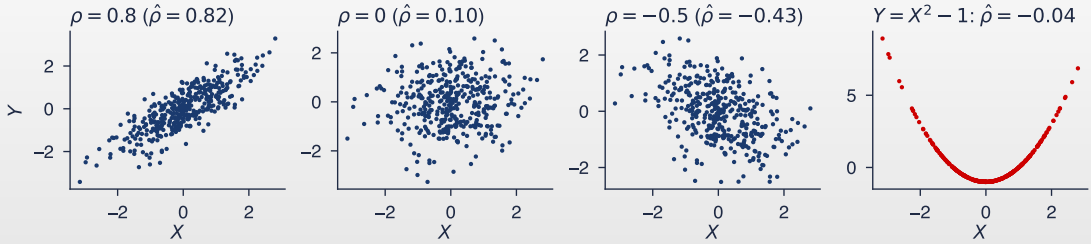
  - punctul cel mai de jos al fiecărei curbe; $\rho = -1$: $\sigma_p = 0$ pentru $a^* = \sigma_2/(\sigma_1 + \sigma_2)$

SFM_ch4_seminar

Noțiuni necesare azi (7/21): independența

- **Independente**: $P(X = x, Y = y) = P(X = x)P(Y = y)$ pentru fiecare celulă a tabelului comun
  - cunoașterea lui X nu schimbă distribuția lui Y
  - independente $\Rightarrow$ necorelate, dar nu și invers
- **Corelația de selecție** pentru n perechi (x_i, y_i): $\hat{\rho} = \dfrac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$
  - $\bar{x}$, $\bar{y}$: mediile de selecție; $\hat{\rho}$ estimează ρ și servește ca test pentru corelația zero
  - dacă datele sînt i.i.d. (independente și identic distribuite), ea se află în banda $\pm 1{,}96/\sqrt{n}$ cu probabilitatea 95%
  - o valoare în afara benzii: un indiciu împotriva independenței

Noțiuni necesare azi (8/21): corelația pe un grafic



- Patru eșantioane simulate a cîte 400 de perechi (X, Y), $X \sim N(0, 1)$
  - primele trei: $Y = \rho X + \sqrt{1 - \rho^2}\, Z$, cu $Z \sim N(0, 1)$ independentă de X, deci $\text{Corr}(X, Y) = \rho$
  - $\hat{\rho}$ diferă puțin de ρ: eroarea de eșantionare
- Dreapta: $Y = X^2 - 1$ este complet determinată de X, totuși $\hat{\rho} \approx 0$
  - $\text{Cov}(X, X^2) = E[X^3] = 0$ pentru un X simetric, cu media 0: corelația surprinde doar dependența liniară

SFM_ch4_seminar

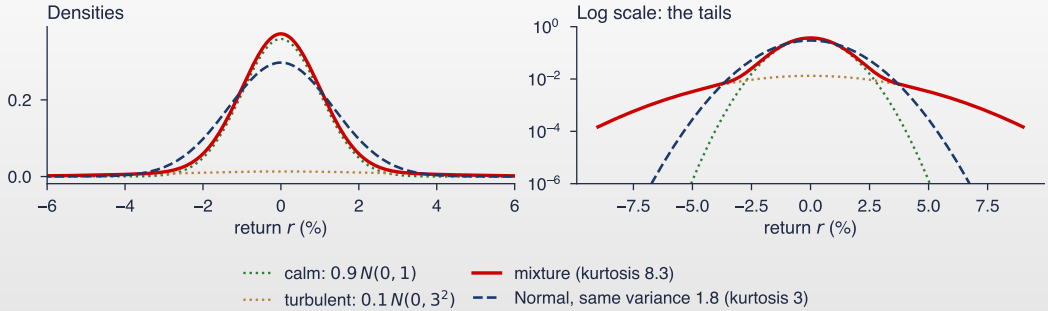
Noțiuni necesare azi (9/21): condiționarea

- $P(A \mid B) = P(A \cap B)/P(B)$
  - probabilitatea lui A după ce aflăm că B a avut loc; $A \cap B$: au loc amîndouă
- $E[Y \mid X]$: media lui Y după ce aflăm X
  - cea mai bună prognoză a lui Y pe baza lui X
  - **proprietatea turnului**: $E[E[Y \mid X]] = E[Y]$, media prognozelor este media generală
- **Legea varianței totale**: $\mathrm{Var}(Y) = E[\mathrm{Var}(Y \mid X)] + \mathrm{Var}(E[Y \mid X])$
  - varianța medie din interiorul grupurilor definite de X, plus varianța mediilor grupurilor

Noțiuni necesare azi (10/21): condiționarea pentru randamente

- $\mathcal{F}_{t-1}$: informația disponibilă pînă ieri (prețuri trecute, știri)
  - $\mu_t = E[r_t \mid \mathcal{F}_{t-1}]$: media condiționată a randamentului de azi
  - $\sigma_t^2 = \mathrm{Var}(r_t \mid \mathcal{F}_{t-1})$: varianța condiționată
  - GARCH modelează tocmai variația în timp a lui σ_t (Capitolul 9)
- Un σ_t variabil face din randamente un amestec de distribuții Normale, cu cozi groase (slide-ul următor)

Noțiuni necesare azi (11/21): un amestec de două regimuri Normale



- Zi calmă cu probabilitatea π: $N(0, \sigma_1^2)$; zi agitată cu probabilitatea $1-\pi$: $N(0, \sigma_2^2)$

$$f(r) = \pi\, \varphi_{\sigma_1}(r) + (1-\pi)\, \varphi_{\sigma_2}(r), \qquad \mathrm{Var}(r) = \pi\sigma_1^2 + (1-\pi)\sigma_2^2$$

  - φ_σ: densitatea $N(0, \sigma^2)$; aici $\pi = 0{,}9$, $\sigma_1 = 1$, $\sigma_2 = 3$: boltirea 8.3, deși fiecare regim este Normal
- Scara logaritmică (dreapta): departe de 0, densitatea amestecului este de multe ori mai mare decît a distribuției Normale cu aceeași varianță: cozi groase

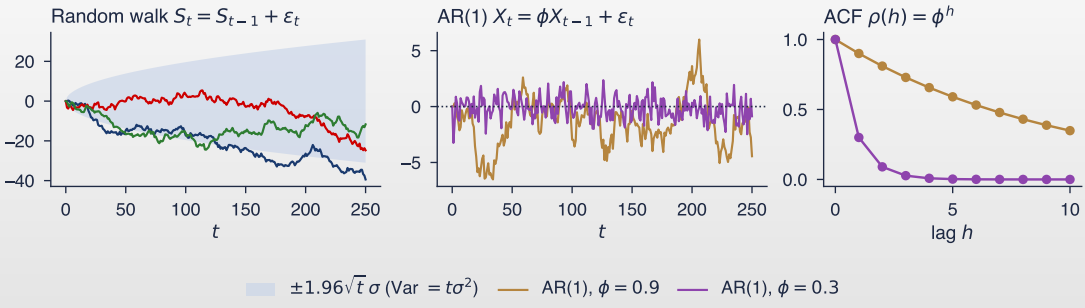
Noțiuni necesare azi (12/21): procese

- **Zgomot alb** ε_t: media 0, varianța σ^2, necorelat în timp
  - zgomot **i.i.d.**: în plus, independent și identic distribuit
- **Mers aleator**: $S_t = S_{t-1} + \varepsilon_t$
  - $E[S_t] = S_0$ (valoarea de pornire), $\mathrm{Var}(S_t) = t\sigma^2$: varianța crește în timp
- **Martingal**: $E[X_{t+1} \mid \mathcal{F}_t] = X_t$
  - un joc echitabil: cea mai bună prognoză pentru mîine este valoarea de azi
  - $\mathcal{F}_t$: informația disponibilă la momentul t
- Proces **staționar** (slab): $E[X_t]$ și $\mathrm{Cov}(X_t, X_{t+h})$ nu depind de t
  - h: lagul dintre două momente; un mers aleator nu este staționar, creșterile lui sînt

Noțiuni necesare azi (13/21): AR(1) și arborele binomial

- **AR(1)**: $X_t = c + \phi X_{t-1} + \varepsilon_t$, staționar dacă $|\phi| < 1$
  - c: o constantă; ϕ: persistența, fracțiunea din valoarea de ieri care trece în ziua de azi
  - media $c/(1-\phi)$, varianța $\sigma^2/(1-\phi^2)$; autocorelația la lagul h: $\rho(h) = \phi^h$
- **Arbore binomial**: la fiecare pas prețul se înmulțește cu u (probabilitatea p) sau cu d (probabilitatea $1-p$)
  - după n pași $S_n = S_0 u^K d^{n-K}$, cu $K \sim B(n,p)$ numărul binomial de creșteri
  - $E[S_n] = S_0(pu + (1-p)d)^n$; joc echitabil pentru $p^* = (1-d)/(u-d)$, soluția ecuației $p^* u + (1-p^*)d = 1$

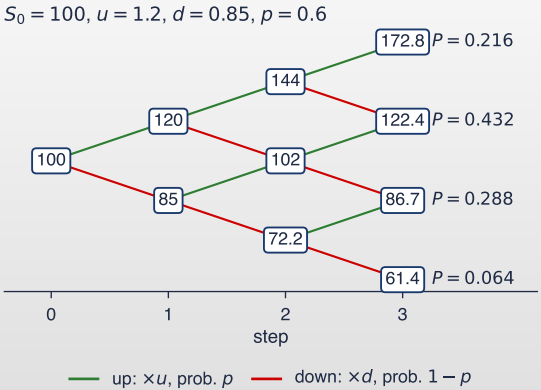
Noțiuni necesare azi (14/21): mersul aleator și AR(1) pe un grafic



- Stînga: trei mersuri aleatoare cu $\sigma = 1$; banda $\pm 1{,}96\sqrt{t}\,\sigma$ se lărgește, deoarece $\mathrm{Var}(S_t) = t\sigma^2$
- Mijloc: traiectoriile AR(1) ($c = 0$) revin mereu la media 0; cu $\phi = 0{,}9$ rămîn mai mult timp departe de ea
- Dreapta: $\rho(h) = \phi^h$ scade geometric; mersul aleator este cazul limită $\phi = 1$, care nu revine niciodată

SFM_ch4_seminar

Noțiuni necesare azi (15/21): arborele binomial pe un grafic



- Numărul de creșteri în n pași: $K \sim B(n, p)$

$$P(K = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, \ldots, n$$

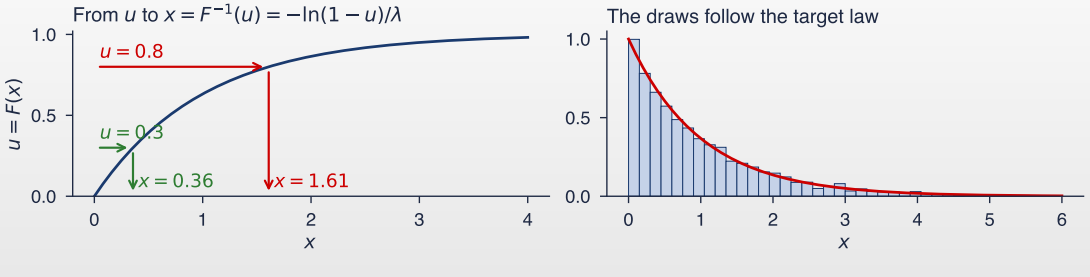
  - $\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$: numărul de traiectorii cu k creșteri
- Arborele se recombină: o creștere urmată de o scădere dă același preț ca o scădere urmată de o creștere, $S_0 ud$
- Graficul: $n = 3$; prețurile finale cu probabilitățile lor, care însumează 1
  - $E[S_3] = S_0(pu + (1-p)d)^3 = 100 \times 1{,}06^3 = 119{,}1$

SFM_ch4_seminar

Noțiuni necesare azi (16/21): numere aleatoare și simulare

- **Transformarea inversă**: dacă $U \sim U(0,1)$ (uniformă pe $[0,1]$), atunci $F^{-1}(U)$ are funcția de repartiție F
  - exponențiala cu rata λ: $-\ln(1-U)/\lambda$; distribuția Normală: $\mu + \sigma\Phi^{-1}(U)$, cu Φ funcția de repartiție $N(0,1)$
- **LCG** (linear congruential generator, generator congruențial liniar): $x_{k+1} = (ax_k + c) \bmod M$, $u_k = x_k/M$
  - a, c, M: numere întregi; $\bmod M$: restul împărțirii la M; perioada: numărul de valori pînă la repetare
- **Bootstrap**: extragem cu întoarcere din randamentele observate
  - transformarea inversă aplicată funcției de repartiție empirice $F_n(x) = \frac{1}{n}\#\{t : r_t \leq x\}$, ponderea observațiilor $\leq x$

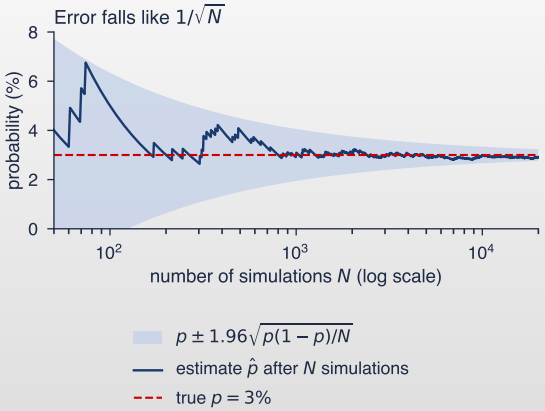
Noțiuni necesare azi (17/21): transformarea inversă pe un grafic



- Stînga: extragem u pe axa verticală și citim $x = F^{-1}(u)$ pe axa orizontală
  - distribuția exponențială cu rata λ: din $u = 1 - e^{-\lambda x}$ rezultă $x = -\ln(1-u)/\lambda$; aici $\lambda = 1$
- Dreapta: 5 000 de astfel de extrageri au densitatea țintă $\lambda e^{-\lambda x}$; un u mai mare dă mereu un x mai mare

SFM_ch4_seminar

Noțiuni necesare azi (18/21): eroarea Monte Carlo și testul binomial



- Estimarea Monte Carlo a unei probabilități p

$$\hat{p} = \frac{1}{N}\sum_{j=1}^{N} 1\{A_j\}, \qquad \mathrm{SE}(\hat{p}) = \sqrt{\frac{p(1-p)}{N}}$$

  - $1\{A_j\} = 1$ dacă evenimentul are loc în simularea j, 0 altfel
- Graficul: $p = 3\%$; banda $\pm 1{,}96\,\mathrm{SE}$ se îngustează ca $1/\sqrt{N}$
  - de 100 de ori mai multe simulări: o eroare de 10 ori mai mică
- **Testul binomial**: un eveniment observat de x ori în m perioade independente
  - dacă probabilitatea p dată de model este corectă, $X \sim B(m, p)$; p-value-ul bilateral $2\min\{P(X \le x), P(X \ge x)\}$
  - un p-value mic (sub 5%): datele contrazic modelul

SFM_ch4_seminar

Noțiuni necesare azi (19/21): procesul Wiener și GBM

- **Procesul Wiener**: $W(0) = 0$, creșteri independente, $W(t) - W(s) \sim N(0, t-s)$
  - varianța unei creșteri este egală cu timpul scurs; $\text{Cov}(W(s), W(t)) = \min(s, t)$
- **GBM** (geometric Brownian motion, mișcarea browniană geometrică)

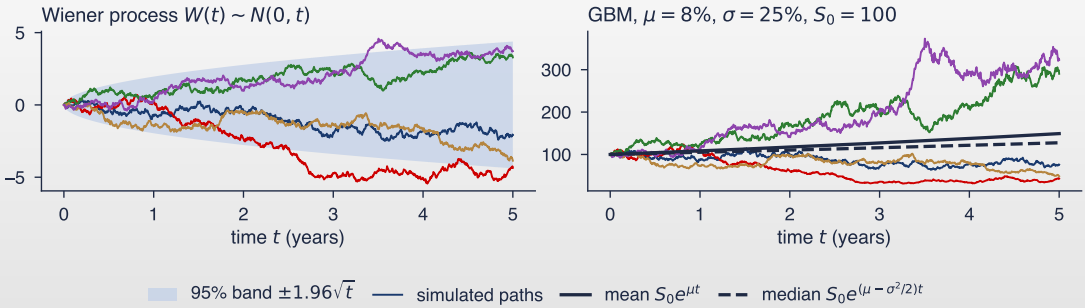
$$S_T = S_0 \exp\left((\mu - \sigma^2/2)T + \sigma W(T)\right)$$

  - μ: tendința anuală (drift); σ: volatilitatea anuală; T: orizontul, în ani; randamente logaritmice zilnice Normale i.i.d.
  - media $S_0 e^{\mu T}$, mediana $S_0 e^{(\mu - \sigma^2/2)T}$
- Prețul logaritmic are distribuția Normală, deci probabilitățile și cuantilele lui S_T rezultă din Φ

$$\ln S_T \sim N\big(\ln S_0 + (\mu - \sigma^2/2)T,\ \sigma^2 T\big), \qquad P(S_T < x) = \Phi\Big(\frac{\ln(x/S_0) - (\mu - \sigma^2/2)T}{\sigma\sqrt{T}}\Big)$$

  - cuantila de ordin α: $S_0 \exp\left((\mu - \sigma^2/2)T + z_\alpha \sigma\sqrt{T}\right)$, cu $z_\alpha = \Phi^{-1}(\alpha)$

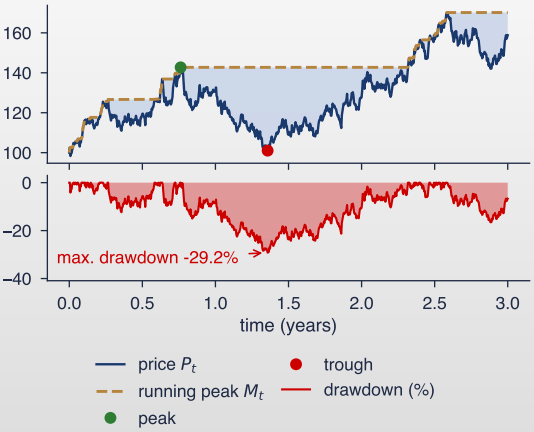
Noțiuni necesare azi (20/21): traiectorii Wiener și traiectorii GBM



- Stînga: cinci traiectorii Wiener; $W(t) \sim N(0, t)$, deci circa 95% dintre traiectorii rămîn în banda $\pm 1{,}96\sqrt{t}$
- Dreapta: aceleași traiectorii transformate în prețuri GBM, $S_0 = 100$, $\mu = 8\%$, $\sigma = 25\%$
  - media $S_0 e^{\mu t}$ este peste mediana $S_0 e^{(\mu - \sigma^2/2)t}$: cîteva traiectorii foarte ridicate trag media în sus

SFM_ch4_seminar

Noțiuni necesare azi (21/21): drawdown-ul



- **Maximul curent**: cel mai mare preț pînă în ziua t

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

- **Drawdown-ul**: scăderea relativă față de maximul curent

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

  - $DD_t = 0$ la fiecare nou maxim

- **Drawdown-ul maxim**: $\mathrm{MDD} = \min_t DD_t$
  - cea mai mare scădere de la un maxim la un minim ulterior
  - graficul: o traiectorie GBM simulată, $\mathrm{MDD} = -29{,}2\%$, de la punctul verde la cel roșu

A1: momentele unui randament discret [Rezolvat]

Cerință

- Randamentul zilnic X (în %) al unei acțiuni ia valorile -3, 0 și 2 cu probabilitățile 0,2, 0,3 și 0,5.
- 1. Calculați $E[X]$, $E[X^2]$, $\mathrm{Var}(X)$ și abaterea standard.
- 2. Scrieți CDF $F(x)$ pentru orice x.
- 3. Calculați $P(X < 0)$ și mediana.
- Raportați: patru momente, CDF și două valori.

Rezolvare

- 1. $E[X] = -0{,}6 + 0 + 1 = 0{,}40$; $E[X^2] = 1{,}8 + 2 = 3{,}80$; $\mathrm{Var} = 3{,}80 - 0{,}40^2 = 3{,}64$; abaterea standard 1,908
- 2. $F(x) = 0$ pentru $x < -3$; 0,2 pe $[-3, 0)$; 0,5 pe $[0, 2)$; 1 pentru $x \geq 2$
- 3. $P(X < 0) = 0{,}2$; mediana $= \inf\{x : F(x) \geq 0{,}5\} = 0$
- Media este pozitivă deși o pierdere este posibilă: riscul și randamentul așteptat sînt întrebări diferite.

înapoi Lista exercițiilor

A2: necorelate, dar dependente [Propus]

Rezolvarea se discută la seminar

Cerință

- X = mișcarea prețului (-1 scădere, 0 stagnare, $+1$ creștere); $Y = 1$ dacă ziua este „mare". Probabilitățile comune: $P(-1,1) = 0{,}25$, $P(0,0) = 0{,}5$, $P(1,1) = 0{,}25$, celelalte celule au probabilitatea 0. Model: A1.
- 1. Calculați distribuțiile marginale ale lui X și Y și mediile lor.
- 2. Calculați $\mathrm{Cov}(X,Y)$ și ρ.
- 3. Decideți dacă X și Y sînt independente, folosind o celulă a tabelului.
- 4. Calculați $\mathrm{Var}(X+Y)$.
- Raportați: două distribuții marginale, ρ, o concluzie și o variantă.

Înapoi Lista exercițiilor

A3: volatilitatea unui portofoliu cu două active [Rezolvat]

Cerință

- Două active cu volatilitățile anuale $\sigma_1 = 20\%$ și $\sigma_2 = 30\%$; ponderile 0,6 și 0,4.
- 1. Calculați volatilitatea portofoliului pentru $\rho = 0{,}3$.
- 2. Repetați pentru $\rho = 0$, $\rho = -0{,}5$ și $\rho = 1$.
- 3. Explicați cînd portofoliul este mai puțin riscant decît media ponderată a celor două volatilități.
- Raportați: patru volatilități și o frază.

Rezolvare

- 1. $\sigma_p = \sqrt{374{,}4} = 19{,}35\%$
  - $\sigma_p^2 = 0{,}36 \times 400 + 0{,}16 \times 900 + 2 \times 0{,}24 \times 0{,}3 \times 600$
- 2. $\rho = 0$: 16,97%; $\rho = -0{,}5$: 12,00%; $\rho = 1$: 24,00%
- 3. Doar $\rho = 1$ dă media ponderată $0{,}6 \times 20 + 0{,}4 \times 30 = 24\%$; orice $\rho < 1$ aduce un cîștig din diversificare.

înapoi Lista exercițiilor

A4: două regimuri și legea variației totale [Propus]

Rezolvarea se discută la seminar

Cerință

- O zi este calmă cu probabilitatea 0,8, $r \mid \text{calm} \sim N(0, 0{,}8^2)$, și agitată în rest, $r \mid \text{agitat} \sim N(0, 2^2)$ (în %). Model: A3.
- 1. Folosiți legea variației totale pentru a calcula $\text{Var}(r)$ și abaterea standard.
- 2. Calculați $E[r^4]$ cu momentul de ordinul 4 al $N(0, \sigma^2)$ (Noțiuni necesare azi (3/21)) în fiecare regim, apoi boltirea $E[r^4]/\text{Var}(r)^2$.
- 3. Explicați de ce amestecul are cozi groase, deși fiecare regim este Normal.
- Raportați: o variantă, o boltire și o frază.

Înapoi Lista exercițiilor

A5: un arbore binomial cu trei pași [Rezolvat]

Cerință

- $S_0 = 100$, $u = 1{,}1$, $d = 0{,}9$, $p = 0{,}55$, trei pași independenți.
- 1. Listați valorile posibile ale lui S_3 și probabilitățile lor.
- 2. Calculați $E[S_3]$ în două moduri: din tabel și din $S_0(pu + (1-p)d)^3$.
- 3. Calculați $P(S_3 < 100)$.
- 4. Găsiți p^* care face din S_t un martingal.
- Raportați: un tabel, două speranțe, o probabilitate, p^*.

Rezolvare

- 1. $K = 0, 1, 2, 3$ creșteri: $72{,}9; 89{,}1; 108{,}9; 133{,}1$ cu $0{,}0911; 0{,}3341; 0{,}4084; 0{,}1664$ $\left(\binom{3}{k} p^k (1-p)^{3-k}\right)$
- 2. $E[S_3] = 100 \times 1{,}01^3 = 103{,}03$ în ambele moduri
- 3. $P(K \leq 1) = 0{,}4253$
- 4. $p^* = (1 - 0{,}9)/(1{,}1 - 0{,}9) = 0{,}50$: cu $p = 0{,}55 > p^*$ prețul are o tendință de creștere

Înapoi Lista exercițiilor

A6: mers aleator, martingal și AR(1) [Propus]

Rezolvarea se discută la seminar

Cerință

- $S_t = S_{t-1} + \varepsilon_t$, ε_t i.i.d. cu media 0 și varianța σ^2; $X_t = 0{,}2 + 0{,}6X_{t-1} + \varepsilon_t$ cu $\sigma = 1$. Model: A5.
- 1. Calculați $E[S_t]$ și $\mathrm{Var}(S_t)$ și arătați că S_t este un martingal.
- 2. Arătați că și $S_t^2 - t\sigma^2$ este un martingal.
- 3. Calculați media, varianța și $\rho(1), \rho(2), \rho(5)$ pentru X_t staționar.
- 4. Calculați $E[X_{t+1} \mid X_t = 2]$ și $\mathrm{Var}(X_{t+1} \mid X_t = 2)$.
- Raportați: două demonstrații, cinci momente, o prognoză.

Înapoi Lista exercițiilor

A7: transformarea inversă pas cu pas [Rezolvat]

Cerință

- Aveți numerele uniforme $u = 0{,}3$ și $u = 0{,}9$ (și $u = 0{,}15, 0{,}45, 0{,}8$ pentru punctul 2).
- 1. Transformați $u = 0{,}3$ și $u = 0{,}9$ în extrageri din distribuția exponențială cu rata $\lambda = 0{,}5$.
- 2. Transformați $u = 0{,}15, 0{,}45, 0{,}8$ în extrageri ale randamentului X din A1.
- 3. Transformați $u = 0{,}01$ într-o extragere din $N(0{,}05, 1{,}2^2)$.
- Raportați: șase valori simulate.

Rezolvare

- 1. $x = -\ln(1-u)/0{,}5$: 0,713 și 4,605
- 2. $F^{-1}(u) = \inf\{x : F(x) \geq u\}$ cu $F = 0{,}2, 0{,}5, 1$: -3, 0 și 2
- 3. $0{,}05 + 1{,}2 \times (-2{,}326) = -2{,}742$: cuantila de 1%
- Un u mare dă o extragere mare: transformarea inversă păstrează ordinea numerelor uniforme.

Înapoi Lista exercițiilor

A8: GBM și procesul Wiener [Propus]

Rezolvarea se discută la seminar

Cerință

- Un indice urmează GBM cu $\mu = 8\%$, $\sigma = 20\%$ pe an, $S_0 = 100$; W este un proces Wiener. Model: A7.
- 1. Calculați media și mediana lui S_1 și ale lui S_5.
- 2. Calculați $P(S_T < 100)$ pentru $T = 1$ și $T = 5$.
- 3. Calculați cuantila de 5% a lui S_1 și a lui S_5.
- 4. Calculați $\mathrm{Var}(W(2) - W(0{,}5))$ și $\mathrm{Cov}(W(1), W(3))$.
- Raportați: opt valori pentru GBM și două momente ale procesului Wiener.

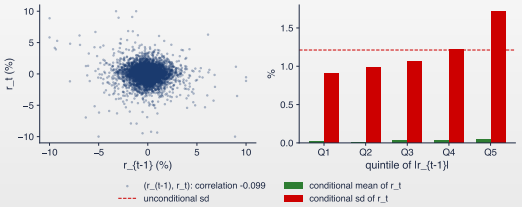
Înapoi Lista exercițiilor

B1: sînt independente randamentele S&P 500? [Rezolvat]

- **Întrebarea**: sînt independente randamentele zilnice ale S&P 500 și ce ne spune ziua de ieri despre cea de azi?
- **Date**: închiderile S&P 500, 2000–2026, randamente logaritmice zilnice în %
- **Cerințe**
  1. Calculați $\mathrm{Corr}(r_t, r_{t-1})$ și $\mathrm{Corr}(r_t^2, r_{t-1}^2)$ și comparați-le cu banda $\pm 1{,}96/\sqrt{n}$.
  2. Împărțiți zilele în chintile ale lui $|r_{t-1}|$ (cinci grupe egale, Q1 cea mai calmă, Q5 cea mai agitată) și calculați media și abaterea standard a lui r_t în fiecare chintilă.
  3. Verificați legea variației totale, cu chintilele ca grupuri.
  4. Interpretare: ce este previzibil pe baza zilei de ieri, media sau varianța randamentului de azi?
- **Raportați**: două corelații, un tabel cu cinci chintile, o descompunere, o frază
- Notebook: secțiunea B1

◂ Înapoi ▸ Lista exercițiilor

B1: rezolvare [Rezolvat]



- $n = 6\,714$, banda $\pm 0{,}024$; $\mathrm{Corr}(r_t, r_{t-1}) = -0{,}099$; $\mathrm{Corr}(r_t^2, r_{t-1}^2) = 0{,}313$
- Mediile condiționate Q1–Q5: 0,021; 0,006; 0,029; 0,027; 0,042 (%)
- Abaterile standard condiționate: 0,91; 0,98; 1,07; 1,21; 1,71; Q5/Q1 $= 1{,}88$
- Varianța totală $1{,}4691 = 1{,}4690$ (în interiorul grupurilor) $+0{,}00013$ (între grupuri, 0,01%)
- Interpretare: varianța este previzibilă, media aproape deloc
  - corelația negativă la lagul 1 provine din zilele de criză cu reveniri mari
  - pătratele sînt corelate, deci randamentele nu sînt independente

SFM_ch4_seminar

B2: aceeași analiză pentru DAX, BET și Bitcoin [Propus]

- **Întrebarea**: arată DAX, BET și Bitcoin același tipar ca S&P 500?
- **Date**: randamente logaritmice zilnice, 2000–2026 (Bitcoin din 2014); model: B1
- **Cerințe**
  1. Calculați ambele corelații la lagul 1 și banda pentru fiecare serie.
  2. Calculați abaterea standard condiționată în Q1 și Q5 și raportul lor.
  3. Calculați ponderea variației explicate de mediile chintilelor.
  4. Interpretare: de ce este BET singura serie cu $\mathrm{Corr}(r_t, r_{t-1})$ clar pozitivă?
- **Raportați**: un tabel și o frază
- Notebook: secțiunea B2

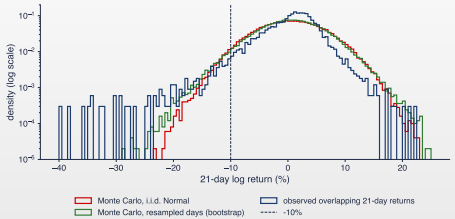
înapoi Lista exercițiilor

B3: probabilitatea unei pierderi lunare prin Monte Carlo [Rezolvat]

- **Întrebarea**: cît de probabilă este o pierdere de peste 10% în 21 de zile pentru S&P 500 și depinde răspunsul de model?
- **Date**: randamentele logaritmice zilnice ale S&P 500, 2000–2026; 100 000 de luni simulate
- **Cerințe**
  1. Calculați probabilitatea exactă în ipoteza unor randamente zilnice Normale i.i.d., cu media și abaterea standard de selecție.
  2. Estimați-o prin Monte Carlo în același model și raportați eroarea standard.
  3. Estimați-o prin reeșantionarea cu întoarcere a 21 de zile observate (transformarea inversă empirică).
  4. Calculați ponderea ferestrelor de 21 de zile suprapuse din date cu o pierdere de peste 10%.
  5. Interpretare: care diferențe se datorează erorii de simulare și care erorii de model?
- **Raportați**: patru probabilități, două erori standard, o frază
- Notebook: secțiunea B3

◂ Înapoi ▸ Lista exercițiilor

B3: rezolvare [Rezolvat]



- Distribuția Normală, exact: 2,922%; Monte Carlo cu distribuția Normală: 2,917% (eroarea standard 0,053%)
  - bootstrap: 3,184% (eroarea standard 0,056%)
- În date: 2,67% din cele 6 694 de ferestre suprapuse (179 de ferestre, concentrate în cîteva episoade)
  - luni care nu se suprapun: 3,45% din 319
- Interpretare
  - Monte Carlo și valoarea exactă diferă cu mai puțin de o eroare standard: eroare de simulare
  - bootstrap-ul și modelul Normal diferă cu circa 3,5 erori standard: eroare de model (cozi zilnice mai groase)
  - datele sînt zgomotoase pentru că ferestrele suprapuse sînt dependente

SFM_ch4_seminar

B4: testarea unui generator de numere aleatoare [Propus]

- **Întrebarea**: poate un generator să treacă testele simple și să fie totuși inutilizabil?
- **Date**: RANDU, $x_{k+1} = 65539\, x_k \bmod 2^{31}$, sămînța 1, 100 000 de numere; NumPy PCG64 cu aceeași sămînță; model: B3
- **Cerințe**
  1. Calculați media, varianța și corelația la lagul 1 pentru ambele șiruri și comparați-le cu $1/2$, $1/12$ și 0.
  2. Numărați valorile distincte ale lui $9u_k - 6u_{k+1} + u_{k+2}$ pentru ambele generatoare.
  3. Găsiți perioada generatoarelor LCG $(a, c, M) = (5, 1, 16)$, $(5, 0, 16)$ și $(4, 1, 16)$ pornite din 7.
  4. Interpretare: de ce nu detectează testele unidimensionale defectul RANDU?
- **Raportați**: un tabel cu șase statistici, două numărători, trei perioade, o frază
- Notebook: secțiunea B4

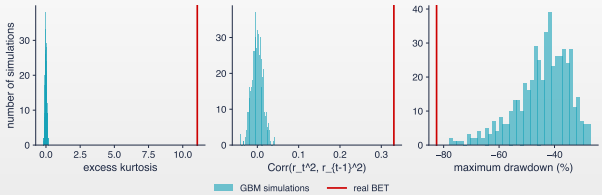
înapoi Lista exercițiilor

B5: este BET o mișcare browniană geometrică? [Rezolvat]

- **Întrebarea**: ce trăsături ale BET poate reproduce un GBM cu aceeași tendință și volatilitate?
- **Date**: închiderile BET, 2000–2026; 500 de traiectorii GBM de aceeași lungime
- **Cerințe**
  1. Calibrați GBM: tendința logaritmică anuală și volatilitatea din randamentele logaritmice zilnice.
  2. Simulați 500 de traiectorii și calculați, pentru fiecare, excesul de boltire, $\mathrm{Corr}(r_t^2, r_{t-1}^2)$ și drawdown-ul maxim.
  3. Plasați valorile reale în distribuțiile simulate (cuantilele de 5% și 95%).
  4. Interpretare: ce fapte stilizate ratează GBM și care indicator de risc este cel mai afectat?
- **Raportați**: doi parametri, trei comparații, o frază
- Notebook: secțiunea B5

◂ Înapoi ▸ Lista exercițiilor

B5: rezolvare [Rezolvat]



- $n = 6\,670$ zile; $\sigma = 22{,}2\%$, tendința logaritmică 16,1% pe an
- Excesul de boltire: real 11,06, intervalul GBM de 90% $[-0{,}10; 0{,}10]$
- Corr(r_t^2, r_{t-1}^2): real 0,330, GBM $[-0{,}022; 0{,}022]$
- Drawdown-ul maxim: real $-82{,}5\%$, mediana GBM $-42{,}7\%$, intervalul $[-62{,}8\%; -31{,}4\%]$
- Interpretare: GBM nu reproduce deloc cozile groase și volatility clustering
  - crahul din 2008 face drawdown-ul real mai adînc decît al oricărei traiectorii simulate
  - o limită de risc calculată cu GBM ar fi fost mult prea largă

SFM_ch4_seminar

B6: crește varianța liniar cu orizontul? [Propus]

- **Întrebarea**: este valabilă relația Var(randament pe h zile) $= h\,$Var(randament zilnic), așa cum implică un mers aleator cu creșteri i.i.d.?
- **Date**: randamentele logaritmice zilnice ale S&P 500, BET, Bitcoin; blocuri disjuncte de $h = 5$, 21 și 63 de zile; model: B5
- **Cerințe**
  1. Calculați raportul Var(h zile)/($h\,$Var(1 zi)) pentru fiecare serie și orizont.
  2. Folosiți eroarea standard aproximativă $\sqrt{2/(m-1)}$ a unui raport de varianțe cu m blocuri pentru a evalua cît de departe este raportul de 1.
  3. Interpretare: ce spune un raport peste 1 sau sub 1 despre autocorelația randamentelor?
- **Raportați**: un tabel 3×3 și două fraze
- Notebook: secțiunea B6

Înapoi Lista exercițiilor

C1: cît de des apare un an nefavorabil? [Propus]

- **Întrebarea**: dă GBM probabilitatea corectă a unui drawdown de peste 20% (sau 40%) în cursul unui an calendaristic?
- **Date**: S&P 500, DAX, BET, Bitcoin, ani calendaristici compleți 2000–2025 (Bitcoin 2015–2025); modele: B3, B5
- **Cerințe**
  1. Calculați drawdown-ul maxim din fiecare an calendaristic și ponderea anilor cu peste 20% și cu peste 40%.
  2. Simulați 10 000 de traiectorii de un an în GBM și în bootstrap-ul i.i.d. al randamentelor zilnice și calculați aceleași două probabilități.
  3. Comparați datele și modelele cu un test binomial.
  4. Propuneți un model care ar putea reproduce ambele probabilități.
  5. Interpretare: ce spune diferența dintre bootstrap și date despre volatility clustering?
- **Raportați**: un tabel, două p-value-uri binomiale, un scurt plan de proiect
- Notebook: secțiunea C1

◂ Înapoi ▸ Lista exercițiilor

C2: verificarea unui răspuns generat de AI [Propus]

- **Un student a cerut unui asistent AI să „rezume faptele de probabilitate despre S&P 500 pentru un raport de risc". Răspunsul:**
- (a) Corr(r_t, r_{t-1}) = −0,099, deci randamentele zilnice sînt independente și varianța de mîine nu poate fi prevăzută.
- (b) Prețul logaritmic este un mers aleator, deci varianța lui este aceeași la fiecare dată și procesul este staționar.
- (c) În GBM cu mu = 8% și sigma = 19%, valoarea mediană a unei investiții de 100 după 10 ani este 100 exp(0,8) = 222,6.
- (d) O probabilitate Monte Carlo de circa 3% din 10.000 de traiectorii are o eroare standard de circa 0,0017.
- (e) Un portofoliu 50/50 din S&P 500 și DAX are volatilitatea zilnică sqrt(0,25 x 1,22^2 + 0,25 x 1,43^2) = 0,94%.
- **Cerințe**
  1. **Pentru fiecare afirmație, precizați dacă este corectă; dacă nu este, formulați afirmația corectă și dați cifra corectă din notebook (secțiunea C2).**
- **Raportați: o listă de cinci verdicte, fiecare cu un rînd de justificare.**

Înapoi Lista exercițiilor

Idei de reținut

- Corelația zero nu înseamnă independență: verificați și pătratele sau valorile absolute
- Varianța condiționată a randamentelor reacționează la știrile de ieri; media condiționată aproape deloc
- Amestecurile de regimuri Normale dau cozi groase
- Monte Carlo: raportați eroarea standard; separați eroarea de simulare de eroarea de model
- GBM reproduce aproximativ traiectoria unui indice, dar nu și cozile, volatility clustering sau cele mai adînci drawdown-uri
- Un răspuns AI este o ciornă: verificați fiecare definiție, fiecare model și fiecare valoare numerică

După seminar

- Cursul 4 dezvoltă fiecare temă de azi: distribuții, dependență, condiționare, Monte Carlo, arbori binomiali, mers aleator, procesul Wiener și GBM
- C1 poate deveni un proiect de echipă: simulări GARCH și cu schimbare de regim, block bootstrap, mai mulți indici
- Lectură: Franke, Härdle & Hafner (2019), cap. 3–5; exerciții rezolvate: Borak, Härdle & López-Cabrera (2013)

Bibliografie

- Borak, S., Härdle, W.K., López-Cabrera, B. (2013). *Statistics of Financial Markets: Exercises and Solutions*, 2nd ed. Springer.
- Cox, J.C., Ross, S.A., Rubinstein, M. (1979). Option pricing: a simplified approach. *Journal of Financial Economics*, 7(3), 229–263.
- Franke, J., Härdle, W.K., Hafner, C.M. (2019). *Statistics of Financial Markets: An Introduction*, 5th ed. Springer.
- Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*. Springer.
- Marsaglia, G. (1968). Random numbers fall mainly in the planes. *Proceedings of the National Academy of Sciences*, 61(1), 25–28.