







Statistica piețelor financiare

Capitolul 6: Selecția modelului și managementul riscului



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)

- **AD**: Anderson–Darling (test) — testul Anderson–Darling
- **AI**: Artificial Intelligence (inteligență artificială)
- **AIC**: Akaike Information Criterion (criteriul informațional Akaike)
- **BET**: Bucharest Exchange Trading (index) — indicele de referință al Bursei de Valori București
- **BIC**: Bayesian Information Criterion (criteriul informațional bayesian)
- **BVB**: Bursa de Valori București
- **COVID-19**: COronaVIrus Disease 2019 (boala provocată de noul coronavirus, identificată în 2019)
- **EDF**: Empirical Distribution Function (funcția de repartiție empirică)
- **EM**: Expectation–Maximisation (algorithm) — algoritmul EM: pasul de medie condiționată și pasul de maximizare
- **EMA**: Exponential Moving Average (medie mobilă exponențială)
- **EODHD**: EOD Historical Data (market-data provider) — furnizor de date de piață

Acronime folosite în acest material (2/2)

- **ES**: Expected Shortfall (pierderea așteptată în coadă)
- **FFT**: Fast Fourier Transform (transformata Fourier rapidă)
- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity (heteroscedasticitate condiționată autoregresivă generalizată)
- **GED**: Generalised Error Distribution (distribuția erorii generalizate)
- **KS**: Kolmogorov–Smirnov (test) — testul Kolmogorov–Smirnov
- **LR**: Likelihood Ratio (raportul de verosimilitate)
- **ML**: Maximum Likelihood (verosimilitate maximă)
- **NIG**: Normal Inverse Gaussian (distribution) — distribuția Normală invers Gaussiană
- **PP**: Probability–Probability (plot) — graficul probabilitate–probabilitate
- **QQ**: Quantile–Quantile (plot) — graficul cuantilă–cuantilă
- **RMA**: Rectangular Moving Average — medie mobilă simplă (dreptunghiulară)

Întrebarea de azi și traseul

- **Întrebarea**: avem șapte distribuții candidate pentru randamentele zilnice; pe care să o folosească un manager de risc și cît de siguri putem fi?
  - Capitolele 2, 3 și 5: distribuția Normală eșuează; Student-t, distribuțiile stabile și teoria valorilor extreme descriu fiecare doar o parte a fenomenului
- **Traseul** capitolului
  - candidații și proprietățile pe care trebuie să le reproducă; recapitulare: metoda verosimilității maxime
  - compararea modelelor: teste ale raportului de verosimilitate, AIC și BIC
  - verificarea unui model: teste de concordanță, grafice PP și QQ
  - scopul decide: ce model estimează corect VaR 1%, în eșantion și în afara lui
  - supraajustarea și incertitudinea de model

Rezultatele învățării

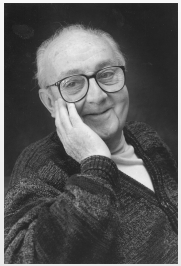
- Scrieți log-verosimilitatea unui model candidat și explicați cum se obțin estimările de verosimilitate maximă
- Testați un model restrîns față de unul mai larg cu testul raportului de verosimilitate și știți cînd testul nu se aplică
- Calculați AIC, BIC, ΔAIC și ponderile Akaike și interpretați-le
- Aplicați și interpretați testele KS, Cramér–von Mises și Anderson–Darling, graficele PP și QQ, și explicați de ce KS este slab în cozi
- Judecați un model după VaR 1% în afara eșantionului, recunoașteți supraajustarea și raportați incertitudinea de model

Bibliografie și instrumente

- Manual: Franke, Härdle & Hafner (2019), cap. 16 (Value at Risk și backtesting) și cap. 18 (statistica riscurilor extreme); exerciții: Borak, Härdle & López-Cabrera (2013)
  - proprietățile distribuției randamentelor: Tsay (2010), cap. 1
- Selecția modelului: Burnham & Anderson (2002); versiunea scurtă Burnham & Anderson (2004)
- Concordanța: Stephens (1974); lucrările originale Anderson & Darling (1952), Lilliefors (1967)
- Quantlet-urile Python ale capitolului: Quantlets/Ch_06
  - adaptate după Quantlet-urile SFESumm și SFEVaRqqplot; fiecare grafic are un link către Quantlet-ul care îl generează
- Notebook-ul cursului: deschideți în Google Colab
- Curs video: Measuring Statistical Risk

„Toate modelele sînt greșite”

- George Box: un statistician caută un model util, nu unul adevărat Box (1976)
  - „toate modelele sînt greșite, dar unele sînt utile” este forma scurtă a argumentului său
- Pentru randamente, „util” depinde de scop
  - evaluarea activelor și alegerea portofoliului: corpul distribuției (zilele obișnuite)
  - managementul riscului: coada stîngă (cele mai slabe zile, 1% din total)
- Acest capitol: instrumentele de comparare a modelelor și obiceiul de a întreba „util pentru ce?”



George E. P. Box (1919–2013)

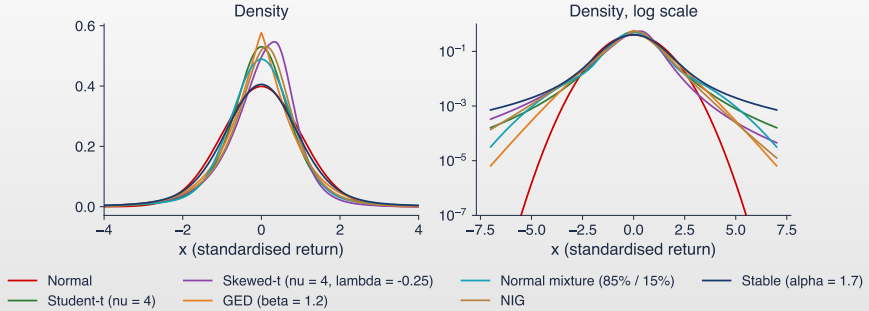
Foto: DavidMCEddy (2011); CC BY-SA 3.0; Wikimedia Commons

Șapte distribuții candidate

Model	k	Cozi	Asimetrie	Referință
Normală	2	subțiri, $e^{-x^2/2}$	nu	Capitolul 2
Student-t	3	putere, $x^{-\nu}$	nu	Capitolul 2
skewed-t	4	putere, $x^{-\nu}$	da (λ)	Hansen (1994)
GED	3	$e^{-\lvert x\rvert^{\beta}}$, groase dacă $\beta < 2$	nu	Nelson (1991)
amestec Normal	5	subțiri, dar o componentă largă	da	Kon (1984)
NIG	4	semigroase, $\lvert x\rvert^{-3/2} e^{-c\lvert x\rvert}$	da	Barndorff-Nielsen (1997)
stabilă	4	putere, $x^{-\alpha}$, $\alpha < 2$	da (β)	Capitolul 3

- k: numărul de parametri; GED: generalised error distribution (distribuția erorii generalizate); NIG: normal inverse Gaussian (Normală invers Gaussiană), din familia hiperbolică generalizată Eberlein & Keller (1995)
- Amestecul Normal: cu probabilitatea w, o zi calmă $N(\mu_1, \sigma_1^2)$, altfel o zi agitată $N(\mu_2, \sigma_2^2)$

Compararea distribuțiilor candidate



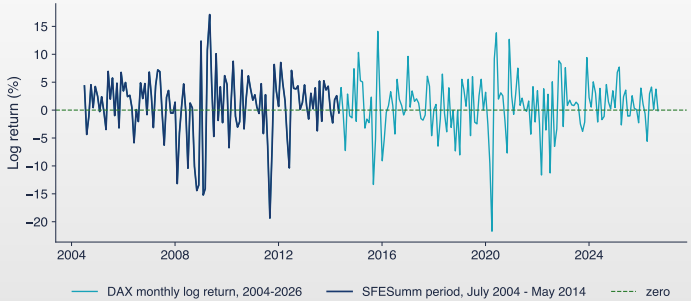
- Toate au media 0 și varianța 1 (distribuția stabilă: același interval intercuartilic ca distribuția Normală); panoul din dreapta folosește scara logaritmică
- În centru arată aproape la fel; în cozi diferă cu cîteva ordine de mărime

SFM_ch6_candidates

Probabilitățile din cozi la variață egală

- $P(X < -4)$, o mișcare de patru abateri standard, pentru fiecare candidat din grafic:
  - Normală $3{,}2 \times 10^{-5}$; GED $9{,}5 \times 10^{-4}$; Student-t $2{,}4 \times 10^{-3}$; amestec Normal $2{,}5 \times 10^{-3}$
  - NIG $3{,}2 \times 10^{-3}$; skewed-t $4{,}5 \times 10^{-3}$; stabilă $8{,}1 \times 10^{-3}$
- Distribuția Student-t atribuie acestei pierderi o probabilitate de circa 76 de ori mai mare decît distribuția Normală
- O măsură de risc precum VaR 1% se află exact în zona în care candidații diferă
- **Întrebare pentru salā**: dacă toate se potrivesc bine în centru, de ce nu îl păstrăm pe cel mai simplu?
- **Răspuns**
  - cel mai simplu model (distribuția Normală) greșește exact în zona care determină cerințele de capital și de marjă

Statisticile de referință pentru modele: SFESumm



- Quantlet-ul SFESumm: randamentele logaritmice lunare DAX între primele zile de tranzacționare, iulie 2004 – mai 2014 ($n = 119$), extinse aici pînă la 1 sep. 2026
- Chiar și randamentele lunare sînt asimetrice și au cozi groase: cifrele sînt pe slide-ul următor

SFM_ch6_summary

Statistici descriptive: DAX lunar și randamente zilnice

	n	min	max	ab. std.	vol. anuală	asim.	boltire
DAX 2004–2014	119	−19,35	17,12	5,80	20,08	−0,83	4,60
DAX 2004–2026	267	−21,70	17,12	5,39	18,66	−0,72	5,06

- Lunar, în %; ab. std.: abaterea standard σ; vol. anuală: $\sigma\sqrt{12}$
  - asimetria $E(X-\mu)^{\mathbf{3}}/\sigma^{\mathbf{3}}$ și boltirea (kurtosis) $E(X-\mu)^{\mathbf{4}}/\sigma^{\mathbf{4}}$, cu μ media; 0 și 3 pentru distribuția Normală
  - testul de normalitate Jarque–Bera (Capitolul 2): p-value $< 0{,}001$
- Randamente logaritmice zilnice din 2000 (Bitcoin din 2014, acțiunile BVB din 2010): boltirea
  - BET 14,1, S&P 500 13,7, DAX 9,1, Bitcoin 15,0, TLV 16,5, SNP 11,9, BRD 10,3
  - asimetria: S&P 500 −0,35, BET −0,65, Bitcoin −0,70

SFM_ch6_summary

Erori în date: două zile din seria Banca Transilvania

- Banca Transilvania (TLV), prețul ajustat de la EODHD: randamente logaritmice de $-15{,}0\%$ pe 30 mai 2016 și $20{,}5\%$ pe 31 mai 2016
  - ajustarea pentru acțiunile gratuite este aplicată cu o zi întîrziere (Capitolul 2): un crah fals urmat de o creștere falsă
- Efectul celor două zile asupra modelului estimat
  - boltirea 20,5 cu ele, 16,5 fără; $\hat{\nu}$ Student-t: 3,33 față de 3,44
  - clasamentul AIC nu se schimbă aici (cel mai bun: Student-t), dar momentele se schimbă
- Regula: înainte de orice selecție de model, comparați cele mai mari randamente cu prețurile de închidere și cu știrile; în restul capitolului, cele două zile sînt eliminate

Verosimilitate și log-verosimilitate

- Datele $r_1, \ldots, r_n$, presupuse i.i.d. (independente și identic distribuite), cu densitatea $f(r; \theta)$
  - θ: vectorul parametrilor modelului (de exemplu μ, σ); f: densitatea modelului
  - **verosimilitatea**: $L(\theta) = \prod_{t=1}^{n} f(r_t; \theta)$, densitatea eșantionului observat, ca funcție de θ
  - **log-verosimilitatea**: $\ell(\theta) = \sum_{t=1}^{n} \ln f(r_t; \theta)$; sumele se manipulează mai ușor decît produsele
- **Estimarea de verosimilitate maximă (ML, maximum likelihood)**: $\hat{\theta} = \arg\max_\theta \ell(\theta)$
  - $\ell(\hat{\theta})$, log-verosimilitatea maximizată, stă la baza tuturor comparațiilor din acest capitol
- Pentru n mare, $\hat{\theta}$ are o distribuție aproximativ Normală în jurul valorii adevărate θ
  - covariația $\approx [-\partial^2 \ell / \partial\theta \, \partial\theta^\top]^{-1}$, inversa informației observate
  - o funcție ℓ puternic curbată (derivata a doua mare) înseamnă o estimare precisă

Exemplu lucrat: modelul Normal, pas cu pas

- $\ell(\mu, \sigma) = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_t (r_t - \mu)^2$
  - anulînd derivatele: $\hat{\mu} = \bar{r}$, $\hat{\sigma}^2 = \frac{1}{n}\sum_t (r_t - \bar{r})^2$ (numitorul n, nu $n-1$)
  - înlocuind în ℓ: $\ell(\hat{\theta}) = -\frac{n}{2}\big[\ln(2\pi\hat{\sigma}^2) + 1\big]$
- S&P 500, randamente logaritmice zilnice în %, 2000–2026: $n = 6\,714$, $\hat{\mu} = 0{,}025$, $\hat{\sigma} = 1{,}213$
  - $\ln(2\pi\hat{\sigma}^2) = 2{,}2240$, deci $\ell(\hat{\theta}) = -\frac{6\,714}{2}(2{,}2240 + 1) = -10\,823$
- Doar modelul Normal are soluție analitică; celelalte se maximizează numeric

Estimarea ML numerică pentru ceilalți candidați (S&P 500)

Model	Estimări	$\ell(\hat{\theta})$
Normală	$\hat{\mu} = 0{,}025$, $\hat{\sigma} = 1{,}213$	$-10\,823$
Student-t	$\hat{\nu} = 2{,}67$, locație $0{,}068$, scală $0{,}703$	$-9\,848$
skewed-t	$\hat{\nu} = 2{,}69$, $\hat{\lambda} = -0{,}067$, $\hat{\mu} = 0{,}024$, $\hat{\sigma} = 1{,}39$	$-9\,838$
GED	$\hat{\beta} = 0{,}88$, locație $0{,}068$, scală $0{,}653$	$-9\,854$
amestec Normal	$\hat{w} = 79\%$: $N(0{,}089; 0{,}71^2)$; 21%: $N(-0{,}223; 2{,}26^2)$	$-9\,925$
NIG	$\hat{a} = 0{,}403$, $\hat{b} = -0{,}048$, locație $0{,}115$, scală $0{,}755$	$-9\,824$
stabilă	$\hat{\alpha} = 1{,}540$, $\hat{\beta} = -0{,}181$, $\hat{\gamma} = 0{,}582$, $\hat{\delta}_0 = 0{,}093$	$-9\,893$

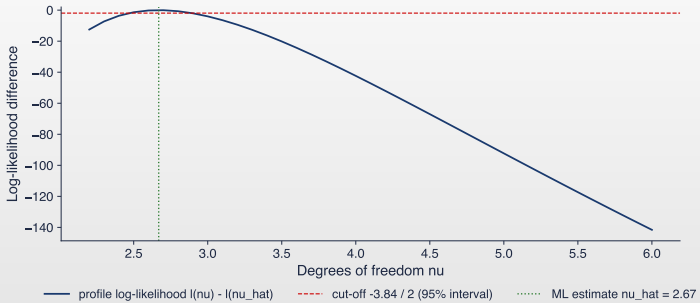
- Algoritmul de optimizare: Nelder–Mead aplicat parametrilor transformați (de exemplu $\ln\sigma$, $\ln(\nu - 2)$); amestecul Normal prin algoritmul EM (expectation–maximisation); densitatea stabilă prin FFT (Capitolul 3)
- Skewed-t: μ și σ sînt media și abaterea standard (parametrizarea lui Hansen)

SFM_ch6_ml_fits

Interpretarea estimărilor

- Student-t $\hat{\nu} = 2{,}67$: cozi de tip $x^{-2{,}67}$; varianța există ($\nu > 2$), boltirea nu există ($\nu < 4$)
  - boltirea de selecție 13,7 estimează o mărime care, conform modelului estimat, este infinită
- Skewed-t $\hat{\lambda} = -0{,}067 < 0$: o coadă stîngă puțin mai lungă (crahuri)
- GED $\hat{\beta} = 0{,}88$: sub 2 (Normală) și chiar sub 1 (Laplace): un vîrf ascuțit
- Amestecul Normal: 79% zile calme cu abaterea standard 0,71%, 21% zile agitate cu 2,26% și medie negativă
- Distribuția stabilă, $\hat{\alpha} = 1{,}540$ (Capitolul 3): cozi mai groase decît orice alt candidat

Precizia lui $\hat{\nu}$: log-verosimilitatea profil



- Pentru fiecare ν, maximizăm ℓ în raport cu locația și scala; la $\nu = 5$, curba se află cu 92 sub maximul său
- Intervalul de încredere de 95%: toate valorile ν pentru care $2[\ell(\hat{\nu}) - \ell(\nu)] \leq 3{,}84$, aici $[2{,}47; 2{,}89]$ în jurul lui $\hat{\nu} = 2{,}67$: grosimea cozilor este estimată precis

SFM_ch6_ml_fits

Probleme numerice în estimarea ML

- **Maxime locale**: optimizatorul se poate opri pe un deal care nu este cel mai înalt
  - remediu: mai multe valori inițiale (amestecurile din acest capitol folosesc între 4 și 12 puncte de pornire)
- **Frontiere**: $\nu \to 2$, o pondere a amestecului $\to 0$, $|\lambda| \to 1$
  - erorile standard și testele obișnuite nu sînt valide la frontieră (secțiunea următoare)
- **Parametrizarea**: aceeași distribuție, parametrizări diferite
  - scala Student-t din scipy nu este abaterea standard: $\mathrm{sd} = \mathrm{scale}\sqrt{\nu/(\nu-2)}$
- **Verificare**: comparați $\ell(\hat{\theta})$ pentru modelele imbricate; un model mai general nu poate avea un maxim mai mic

Modele imbricate

- Modelul 0 este **imbricat** în modelul 1 dacă modelul 0 este modelul 1 cu unii parametri fixați
  - Normală în GED: $\beta = 2$
  - Student-t în skewed-t: $\lambda = 0$
  - Normală în Student-t: $\nu = \infty$, adică $1/\nu = 0$, la **marginea** spațiului parametrilor
- Neimbricate: Student-t și GED, Student-t și stabilă, skewed-t și NIG
  - niciunul nu este caz particular al celuilalt: sînt necesare alte instrumente (testul Vuong, AIC)
- Modele imbricate: întotdeauna $\ell_1(\hat{\theta}_1) \geq \ell_0(\hat{\theta}_0)$; întrebarea este dacă cîștigul depășește ce s-ar obține din întîmplare

Testul raportului de verosimilitate

- H_0: modelul 0 (restrîns) este adevărat; H_1: modelul 1 (mai general)
  - $LR = 2\left[\ell_1(\hat{\theta}_1) - \ell_0(\hat{\theta}_0)\right] \geq 0$: de două ori cîștigul de log-verosimilitate maximizată
  - ℓ_0, ℓ_1: log-verosimilitățile celor două modele; k_0, k_1: numerele lor de parametri
- **Teorema lui Wilks** (Wilks (1938)): sub H_0, $LR \approx \chi^2(d)$, cu $d = k_1 - k_0$ restricții
  - $\chi^2(d)$: distribuția hi-pătrat cu d grade de libertate
  - valabilă dacă valorile fixate se află în interiorul spațiului parametrilor
- Decizia la 5%: respingem H_0 dacă $LR > \chi^2_{0,95}(d)$
  - $\chi^2_{0,95}(1) = 3{,}84$; $\chi^2_{0,95}(2) = 5{,}99$
- Intuiția: de două ori logaritmul raportului dintre verosimilitățile datelor sub cele două modele

Exemplu lucrat: testarea parametrului de asimetrie

- S&P 500: ℓ(Student-t) $= -9\,848$, ℓ(skewed-t) $= -9\,838$
  - $LR = 2(-9\,838 - (-9\,848)) = 19{,}63 > 3{,}84$; $p < 0{,}001$: respingem $\lambda = 0$
- Același test pentru alte serii
  - DAX: $LR = 20{,}92$, $p < 0{,}001$: asimetria este necesară
  - BET: $LR = 0{,}35$, $p = 0{,}555$; Bitcoin: $LR = 0{,}42$, $p = 0{,}515$: nicio dovadă de asimetrie
- Normală în GED ($\beta = 2$): LR între 1 431 (Bitcoin) și 2 298 (BET); modelul Normal este respins peste tot
- Interpretare: cei doi indici dezvoltați au o coadă stîngă mai lungă; BET și Bitcoin nu

SFM_ch6_lr_aic

Un caz de frontieră: Normală în Student-t

- H_0: $1/\nu = 0$ este pe marginea spațiului $1/\nu \geq 0$: teorema lui Wilks nu se aplică
  - Self & Liang (1987): distribuția statisticii sub H_0 este un amestec 50:50 de $\chi^2(0)$ (o masă de probabilitate concentrată în 0) și $\chi^2(1)$
  - valoarea critică la 5%: $\chi^2_{0,90}(1) = 2{,}71$ în loc de 3,84; p-value-ul naiv este dublul celui corect
- Date reale: $LR = 1\,949$ (S&P 500), $1\,337$ (Bitcoin): modelul Normal este respins cu oricare valoare critică
- Aceeași problemă apare pentru o pondere a amestecului $w = 0$ și pentru parametri de varianță egali cu 0
- **Ce credeți?** Contează frontiera cînd LR este de ordinul miilor?
- **Răspuns**
  - nu pentru decizia de aici; contează însă în eșantioane mici și pentru valori LR la limită, unde înjumătățește p-value-ul

Modele neimbricate: testul Vuong

	Student-t față de GED	Student-t față de stabilă	skewed-t față de NIG
BET	3,70 ($<$ 0,001)	9,33 ($<$ 0,001)	$-$1,34 (0,179)
S&P 500	0,40 (0,686)	6,52 ($<$ 0,001)	$-$2,20 (0,028)
DAX	$-$0,23 (0,817)	7,66 ($<$ 0,001)	$-$2,97 (0,003)
Bitcoin	$-$3,80 ($<$ 0,001)	16,47 ($<$ 0,001)	$-$6,76 ($<$ 0,001)

- Vuong (1989): $m_t = \ln f_{\mathbf{1}}(r_t) - \ln f_{\mathbf{2}}(r_t)$, $V = \sqrt{n}\,\bar{m}/s_m \approx N(0,1)$ dacă ambele modele sînt la fel de aproape de adevăr
  - $f_{\mathbf{1}}, f_{\mathbf{2}}$: cele două densități estimate; m_t: cu cît explică modelul 1 mai bine ziua t; $\bar{m}$, s_m: media și abaterea standard a valorilor m_t
  - $V > 1{,}96$: primul model este mai aproape; $V < -1{,}96$: al doilea; p-value-ul este în paranteză
- Student-t este preferat distribuției stabile pe toate piețele; NIG este preferat distribuției skewed-t pentru S&P 500 și DAX; comparația Student-t–GED diferă de la o piață la alta

SFM_ch6_lr_aic

Abordarea lui Akaike: divergența Kullback–Leibler

- Hirotugu Akaike și-a pus întrebarea: ce model va descrie cel mai bine date **noi**?
  - distanța de la distribuția adevărată g la un model f: divergența Kullback–Leibler Kullback & Leibler (1951)
  - $KL(g, f) = E_g[\ln g(R) - \ln f(R)] \geq 0$, zero doar dacă $f = g$
  - R: un randament nou, generat de distribuția adevărată g; E_g: media calculată sub g
- $\ell(\hat{\theta})$ este prea optimistă: aceleași date servesc și la estimarea lui $\hat{\theta}$, și la evaluarea lui
  - supraestimarea este, în medie, de circa k, numărul de parametri Akaike (1974)
- Rezultatul: o penalizare de 2 pentru fiecare parametru corectează supraestimarea; criteriul obținut, AIC, este pe slide-ul următor



Hirotugu Akaike (1927–2009)

Foto: The Institute of Statistical Mathematics; CC BY-SA 4.0; Wikimedia Commons

AIC și BIC

- **AIC** (Akaike information criterion, criteriul informațional Akaike): $\mathrm{AIC} = -2\ell(\hat{\theta}) + 2k$
  - scopul: cel mai bun model predictiv; asimptotic echivalent cu validarea încrucișată leave-one-out Stone (1977)
  - eșantioane mici: $\mathrm{AICc} = \mathrm{AIC} + 2k(k+1)/(n-k-1)$ Hurvich & Tsai (1989)
- **BIC** (Bayesian information criterion, criteriul informațional bayesian) Schwarz (1978): $\mathrm{BIC} = -2\ell(\hat{\theta}) + k \ln n$
  - scopul: modelul adevărat, dacă este printre candidați (consistență)
  - penalizarea pe parametru $\ln n$: 8,81 pentru $n = 6\,714$, de peste patru ori mai mare decît penalizarea AIC, egală cu 2
- Pentru ambele criterii, valoarea mai mică este preferată; doar diferențele dintre modele estimate pe **aceleași date** sînt interpretabile

ΔAIC și ponderile Akaike

- $\Delta_i = \mathrm{AIC}_i - \min_j \mathrm{AIC}_j$; regulile practice din Burnham & Anderson (2004):
  - $\Delta \leq 2$: sprijin substanțial; $4 \leq \Delta \leq 7$: considerabil mai puțin; $\Delta > 10$: practic niciunul
- **Ponderea Akaike**: $w_i = \dfrac{e^{-\Delta_i/2}}{\sum_j e^{-\Delta_j/2}}$
  - ponderea dovezilor în favoarea modelului i dintre candidați; suma ponderilor este 1
  - două modele cu $\Delta = 0$ și $\Delta = 2$: $w = 1/(1 + e^{-1}) \approx 0{,}73$ și $0{,}27$
- **Ce credeți?** Este $\Delta = 2$ o diferență „semnificativă”?
- **Răspuns**
  - nu: AIC nu este un test; o diferență de 2 înseamnă că ambele modele rămîn plauzibile (ponderile de mai sus)

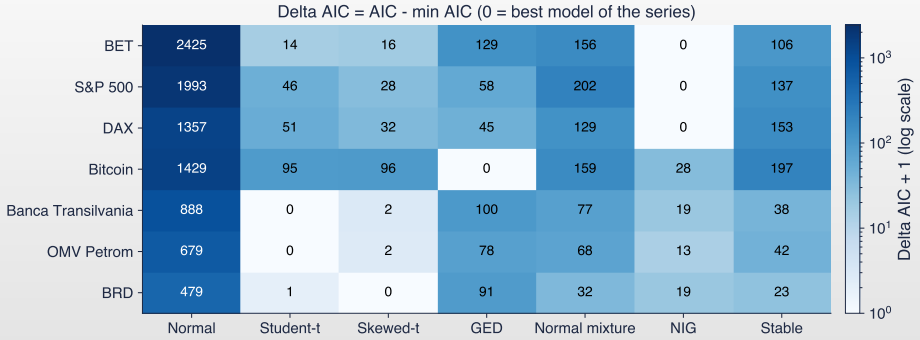
Exemplu lucrat: șapte modele pentru S&P 500

	k	$\ell(\hat\theta)$	AIC	BIC	ΔAIC	ΔBIC	w
Normală	2	−10 823	21 650	21 663	1993,2	1979,6	0,000
Student-t	3	−9 848	19 702	19 723	45,7	38,9	0,000
skewed-t	4	−9 838	19 685	19 712	28,1	28,1	0,000
GED	3	−9 854	19 714	19 735	57,9	51,1	0,000
amestec Normal	5	−9 925	19 859	19 893	202,5	209,3	0,000
NIG	4	−9 824	19 657	19 684	0,0	0,0	1,000
stabilă	4	−9 893	19 794	19 821	137,3	137,3	0,000

- Exemplu: NIG, $\mathrm{AIC} = -2(-9\,824) + 2 \cdot 4 = 19\,657$
- NIG câștigă pe ambele criterii; următorul model, skewed-t, are $\Delta = 28{,}1$, deci $e^{-\Delta/2} \approx 8{,}1 \times 10^{-7}$: ponderea $w \approx 1$ pentru NIG
- Modelul Normal: ΔAIC de 1 993; distribuția stabilă: 137

SFM_ch6_lr_aic

ΔAIC pentru șapte serii



- Fiecare rînd: o serie; 0 marchează cel mai bun model; celule mai închise: modele mai slabe (scară logaritmică a culorii)

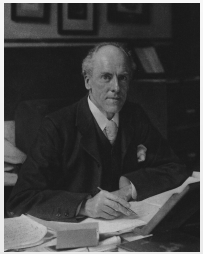
SFM_ch6_lr_aic

Interpretarea hărții ΔAIC

- Niciun cîștigător unic: cel mai bun model depinde de piață
  - indicii (BET, S&P 500, DAX): NIG; Bitcoin: GED; acțiunile BVB: Student-t (TLV, SNP) și skewed-t (BRD)
- Concluzii robuste
  - modelul Normal este ultimul peste tot, cu ΔAIC între 479 (BRD) și 2 425 (BET)
  - distribuția stabilă (Δ de la 23 la 197) și amestecul Normal nu sînt niciodată competitive
- **Întrebare pentru sală**: pentru BRD, AIC preferă skewed-t (Δ al Student-t: 1,0); ce model preferă BIC?
- **Răspuns**
  - Student-t: un parametru în minus reduce BIC cu $\ln n \approx 8{,}2$, mai mult decît cîștigul de ajustare; cu $\Delta \approx 1$, cele două sînt practic la egalitate

Testul χ^2 al lui Pearson și funcția de repartiție empirică

- Pearson (1900): primul test de concordanță, testul χ^2
  - grupăm datele în clase și comparăm frecvențele observate cu cele așteptate
  - rezultatul depinde de clase, iar cozile conțin puține observații
- **Funcția de repartiție empirică** (EDF, empirical distribution function): $F_n(x) = \frac{1}{n}\#\{t : r_t \leq x\}$
  - fără clase: comparăm F_n cu funcția de repartiție a modelului F
- Trei teste EDF: Kolmogorov–Smirnov, Cramér–von Mises, Anderson–Darling Stephens (1974)

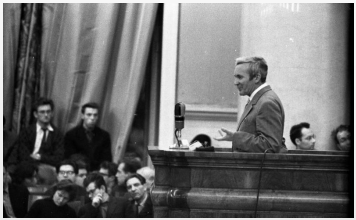


Karl Pearson (1857–1936)

Fotogravură (1910); domeniu public; Wikimedia Commons

Testul Kolmogorov–Smirnov

- $D_n = \sup_x |F_n(x) - F(x)|$, cea mai mare distanță pe verticală dintre EDF și funcția de repartiție a modelului (sup: cea mai mare valoare pe toate x)
  - $r_{(i)}$: al i-lea cel mai mic randament; $u_{(i)} = F(r_{(i)})$: probabilitatea lui în model
  - formula de calcul: $D_n = \max_i \max\{i/n - u_{(i)},\ u_{(i)} - (i-1)/n\}$
- F complet cunoscută: distribuția lui D_n nu depinde de F (Kolmogorov, 1933)
  - n mare: respingem la 5% dacă $D_n > 1{,}358/\sqrt{n}$; tabele pentru n mic: Smirnov (1948)
  - $n = 6\,714$: valoarea critică 0,017
- Andrei Kolmogorov este și autorul axiomaticii probabilităților (Capitolul 4)



Andrei Kolmogorov (1903–1987) la curs, 1963

Foto: Vsevolod Tarasevich (1963–1964); CC BY 4.0; Wikimedia Commons

Exemplu lucrat: KS pe cinci randamente

- Randamente (%): $-2{,}1;\ -0{,}4;\ 0{,}3;\ 0{,}9;\ 1{,}6$; H_0: $N(0,1)$, complet specificată
  - $u_{(i)} = \Phi(r_{(i)})$: 0,018; 0,345; 0,618; 0,816; 0,945 (Φ: funcția de repartiție Normală standard)
  - i/n: 0,2; 0,4; 0,6; 0,8; 1,0; cea mai mare distanță este $u_{(3)} - 2/5 = 0{,}218$
- $D_5 = 0{,}218 < 0{,}563$, valoarea critică exactă la 5% pentru $n = 5$: nu respingem
- Cu cinci observații, aproape orice distribuție trece testul: în eșantioane mici, puterea testului este redusă

SFM_ch6_gof_tests

Parametri estimați: problema Lilliefors

- În practică, F are parametri estimați: $F(x;\hat{\theta})$ este estimată pe aceleași date
  - estimarea apropie F de F_n: D_n este mai mic decît pentru o F cunoscută
  - valorile critice obișnuite sînt prea mari: testul aproape că nu mai respinge
- Lilliefors (1967): valori critice obținute prin simulare pentru distribuția Normală cu $\hat{\mu}$, $\hat{\sigma}$ (nivelul 5%)
  - simularea din acest capitol: $n = 5$: 0,351 în loc de 0,563; $n = 1000$: 0,029 în loc de 0,043
- Orice model: **bootstrap parametric** Stute et al. (1993)
  - simulăm n extrageri din $F(\cdot;\hat{\theta})$, reestimăm parametrii, recalculăm statistica; repetăm de B ori
  - p-value $= (1 + \#\{\text{statistici bootstrap} \geq \text{cea observată}\})/(B+1)$: proporția statisticilor simulate cel puțin egale cu cea observată

Cramér–von Mises și Anderson–Darling

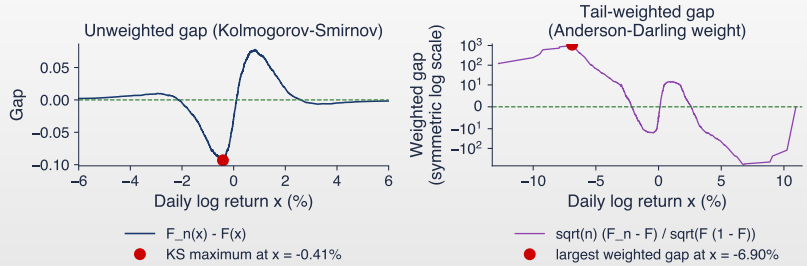
- Ambele integrează pătratul distanței $[F_n(x) - F(x)]^2$ pe toate valorile x în loc să ia maximul
  - **Cramér–von Mises**: $W^2 = n \int [F_n - F]^2 \, dF = \frac{1}{12n} + \sum_i \left(u_{(i)} - \frac{2i-1}{2n}\right)^2$ Cramér (1928)
  - **Anderson–Darling**: $A^2 = n \int \frac{[F_n - F]^2}{F(1-F)} \, dF$ Anderson & Darling (1952)
- Formula de calcul Anderson & Darling (1954):
  $A^2 = -n - \frac{1}{n} \sum_i (2i-1) \left[\ln u_{(i)} + \ln(1 - u_{(n+1-i)}) \right]$
  - $u_{(i)} = F(r_{(i)})$, ca la KS; valori mari ale lui W^2 sau A^2: o ajustare slabă
  - ponderea $1/[F(1-F)]$ este foarte mare acolo unde F este aproape de 0 sau 1, adică în cozi



Harald Cramér (1893–1985)

Foto: fotograf necunoscut (1951); domeniu public; Wikimedia Commons

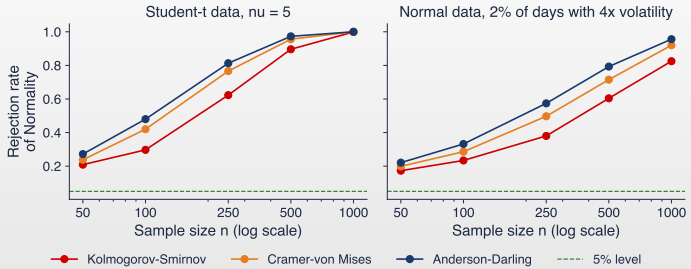
Puterea redusă a testului KS în cozi



- S&P 500 comparat cu distribuția Normală estimată. Stînga: distanța KS are maximul la $x = -0{,}41\%$, aproape de centru ($F = 36\%$), cu $D_n = 0{,}093$
- În cozi, F_n și F sînt amîndouă aproape de 0 sau 1, deci distanța brută este mică chiar dacă modelul greșește mult
  - cu ponderea Anderson–Darling (dreapta), distanța maximă apare la $x = -6{,}9\%$

SFM_ch6_gof_tests

Puterea celor trei teste



- Proporția din 1000 de eșantioane simulate în care normalitatea (parametri estimați) este respinsă la 5%; puterea = probabilitatea de a respinge un H_0 fals
- Date Student-t, $n = 250$: KS 62%, Cramér–von Mises 77%, Anderson–Darling 81%
  - abateri doar în cozi ($n = 250$): 38%, 50%, 57%

SFM_ch6_gof_tests

Teste EDF pe S&P 500

	KS	CvM	AD
Normală	0,093	22,84	127,35
Student-t	0,018	0,57	4,69
skewed-t	0,017	0,41	2,25
GED	0,019	0,38	3,94
amestec Normal	0,025	1,09	5,80
NIG	0,010	0,13	0,88
stabilă	0,022	0,81	4,61

- Valorile mai mici ale statisticilor indică o ajustare mai bună

- Bootstrap parametric, $B = 199$, valori critice (cuantila de 95%)
  - Student-t: KS 0,009, AD 0,71
  - skewed-t: KS 0,008, AD 0,35
- Normală, Student-t, skewed-t: p bootstrap $= 0{,}005$, cel mai mic posibil: toate respinse
- Student-t, p-value-ul KS naiv (ca și cum θ ar fi cunoscut): 0,021

SFM_ch6_gof_tests

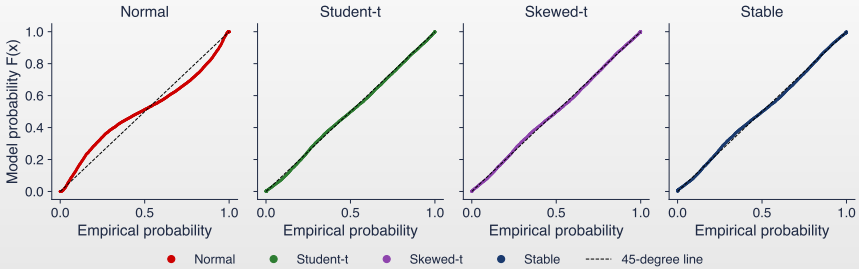
Semnificație statistică și semnificație practică

- Cu $n = 6\,714$, toate modelele sînt respinse: testele detectează abateri minuscule
  - în plus, randamentele reale nu sînt i.i.d. (volatility clustering), deci nicio distribuție i.i.d. nu poate trece testul
- Folosiți statisticile pentru a **ordona** modelele
  - coloana AD din tabelul anterior ordonează modelele aproape ca AIC: NIG pe primul loc, skewed-t pe al doilea, modelul Normal mult în urmă (AD plasează GED și distribuția stabilă înaintea distribuției Student-t)
- **Întrebare pentru sală**: un test KS pe 50 de randamente zilnice nu respinge distribuția Normală; este atunci distribuția Normală potrivită pentru VaR 1%?
- **Răspuns**
  - nu: cu $n = 50$, testul respinge rar chiar și date Student-t (în 30% din cazuri la $n = 100$), iar KS aproape ignoră coada

Definițiile graficelor PP și QQ

- **Grafic PP** (probabilitate–probabilitate): punctele $\left(p_i, F(r_{(i)})\right)$ cu $p_i = (i - 0{,}5)/n$
  - ambele axe în $[0, 1]$: se văd diferențele din centru; cozile sînt înghesuite în colțuri
  - echivalentul grafic al testului KS: cea mai mare distanță pe verticală față de diagonală este apropiată de D_n
- **Grafic QQ** (cuantilă–cuantilă): punctele $\left(F^{-1}(p_i), r_{(i)}\right)$
  - în unități de randament: cozile sînt dilatate, iar fiecare zi extremă este vizibilă
  - capetele se îndepărtează de dreaptă: cozile modelului sînt prea scurte; capetele aplatizate: cozile modelului sînt prea lungi
- În managementul riscului se folosesc graficele QQ; graficele PP și testul KS au același punct orb

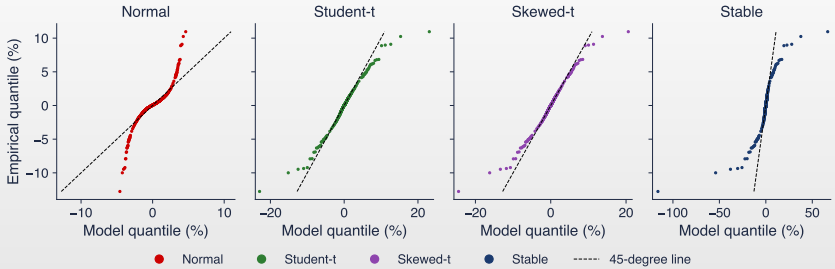
Grafice PP: S&P 500



- Normală: o formă de S în centru (prea puține randamente mici, prea multe de mărime medie), cea mai mare distanță 0,093
- Student-t, skewed-t, stabilă: pe diagonală; într-un grafic PP nu se pot deosebi

SFM_ch6_pp_qq

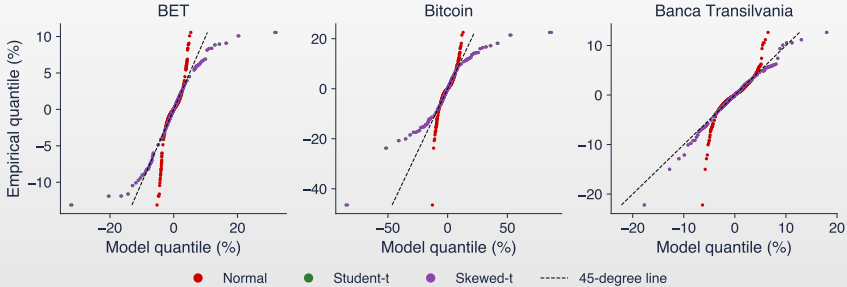
Grafice QQ: S&P 500



- Cel mai mic randament zilnic a fost $-12{,}8\%$; cuantila modelului la aceeași probabilitate:
  - Normală $-4{,}6\%$, Student-t $-22{,}9\%$, skewed-t $-24{,}5\%$, stabilă $-116{,}1\%$
- Modelele care păreau identice în graficul PP diferă puternic în extremitatea cozii

SFM_ch6_pp_qq

Grafice QQ pentru BET, Bitcoin și Banca Transilvania



- Normală (roșu): capete abrupte pentru toate cele trei serii; Student-t și skewed-t urmăresc datele pînă departe în cozi
- Ultimele cîteva puncte sînt zile izolate (cele mai mari crahuri): acolo cozile estimate diferă cel mai mult, iar incertitudinea este cea mai mare

SFM_ch6_pp_qq

VaR și ES

- Randamentul X, nivelul α (aici $\alpha = 1\%$ sau 2,5%, probabilitatea cozii)
  - **VaR** (Value at Risk, valoarea expusă la risc): $\mathrm{VaR}_\alpha = -q_\alpha(X)$, pierderea depășită cu probabilitatea α
  - **ES** (expected shortfall, pierderea așteptată în coadă): $\mathrm{ES}_\alpha = -\frac{1}{\alpha}\int_0^\alpha q_u\,du = -E[X \mid X \le q_\alpha]$, pierderea medie din coadă
- Modelul Normal: $\mathrm{VaR}_\alpha = -(\mu + \sigma z_\alpha)$, $z_{0,01} = -2{,}326$
  - S&P 500: $-(0{,}025 - 2{,}326 \times 1{,}213) = 2{,}80\%$, față de VaR 1% empiric de 3,45%
- Regulile Basel: ES 2,5% pentru riscul de piață, VaR 1% pentru backtesting (Capitolul 10)

VaR 1% și ES 2,5% pentru fiecare model (în eșantion)

VaR 1% (%)

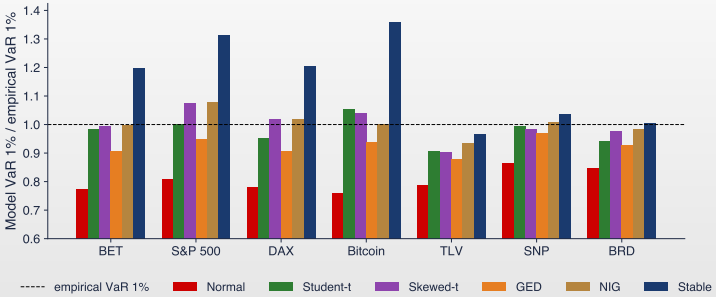
	date	Normală	Student-t	skewed-t	GED	amestec	NIG	stabilă
BET	4,13	3,19	4,06	4,10	3,74	4,53	4,11	4,94
S&P 500	3,45	2,80	3,46	3,71	3,27	3,97	3,72	4,54
DAX	4,16	3,25	3,96	4,24	3,77	4,35	4,24	5,01
Bitcoin	10,51	7,99	11,06	10,91	9,85	10,40	10,52	14,26
TLV	5,10	4,01	4,63	4,61	4,49	5,41	4,76	4,92
SNP	4,29	3,70	4,26	4,22	4,16	4,70	4,32	4,43
BRD	4,50	3,81	4,24	4,39	4,18	4,74	4,43	4,51

ES 2,5% (%)

	date	Normală	Student-t	skewed-t	GED	amestec	NIG	stabilă
BET	4,57	3,21	4,76	4,81	3,85	4,53	4,31	7,58
S&P 500	3,76	2,81	3,95	4,23	3,36	3,97	3,89	6,70
Bitcoin	10,84	8,03	13,37	13,18	10,17	10,44	11,06	22,87

SFM_ch6_tail_var

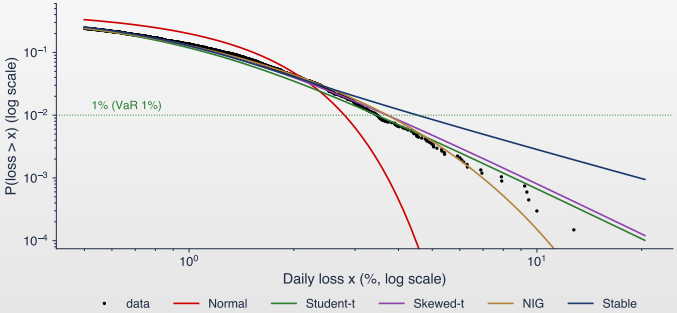
VaR 1% al modelelor raportat la date



- Normală: prea mic cu 14–24% pe toate seriile; stabilă: prea mare cu pînă la 36%
- Student-t, skewed-t și NIG se abat cu cel mult 10% de la VaR 1% empiric; GED este puțin prea mic

SFM_ch6_tail_var

Coada stîngă pe axe log-log (S&P 500)



- Punctele: proporția zilelor cu o pierdere peste x; liniile: aceeași probabilitate pentru fiecare model estimat
- La 1%, toate modelele cu cozi groase sînt apropiate; pentru pierderi de peste 5%, curba distribuției stabile este mult prea sus, iar cea Normală scade abrupt

SFM_ch6_tail_var

Numărarea depășirilor: testul Kupiec

	așteptat	Normală	Student-t	skewed-t	GED	NIG	stabilă
BET	67	138	71	69	91	68	43
S&P 500	67	131	67	52	80	52	30
DAX	68	145	78	66	91	65	39
Bitcoin	44	86	34	36	56	44	16

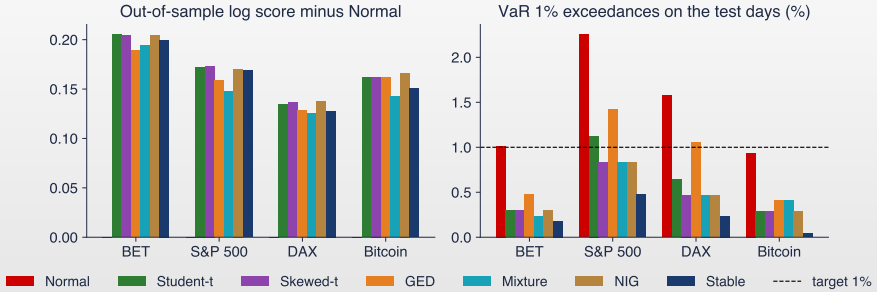
- Depășire: o zi cu $r_t < -\mathrm{VaR}_{1\%}$; pentru un model corect, numărul lor este $\mathrm{Binomial}(n; 0{,}01)$
- Kupiec (1995): $LR_{uc} = -2 \ln \dfrac{(1-\alpha)^{n-x}\alpha^x}{(1-x/n)^{n-x}(x/n)^x} \approx \chi^2(1)$, x = numărul de depășiri
  - un test LR care compară rata țintă α cu rata observată x/n; LR_{uc} mare: număr greșit de depășiri
- S&P 500: Student-t, 67 de depășiri față de 67 așteptate ($p = 0{,}986$)
  - NIG, cîștigătorul AIC: 52 ($p = 0{,}053$)
- Cel mai bun model pentru întreaga densitate nu este întotdeauna cel mai bun pentru o cuantilă; există scoruri axate pe coadă Diks, Panchenko & van Dijk (2011)

SFM_ch6_tail_var

Rolul validării în afara eșantionului

- În eșantion, un model mai general se ajustează cel puțin la fel de bine: ℓ nu scade niciodată cînd adăugăm parametri
  - **supraajustare** (overfitting): parametrii suplimentari modelează zgomotul eșantionului de estimare, nu trăsături ale randamentelor viitoare
- Verificarea onestă: estimăm pînă la 31 decembrie 2019 și evaluăm pe 2020–2026 (crahul COVID-19, șocul inflației din 2022, taxele vamale din 2025)
  - **scorul logaritmic**: media lui $\ln f(r_t; \hat{\theta})$ pe zilele de test; valoarea mai mare este preferată; este o regulă de scor proprie (proper scoring rule) Gneiting & Raftery (2007)
  - **depășirile** VaR 1% (Kupiec) și **pierderea pinball** $\frac{1}{n}\sum_t (\alpha - 1\{r_t < q\})(r_t - q)$ a cuantilei de 1% q
  - pinball: $1\{r_t < q\} = 1$ într-o zi sub q, 0 altfel; penalizează asimetric abaterile, valoarea mai mică este preferată
- Un scor este **propriu** dacă distribuția adevărată obține cel mai bun scor așteptat: prognozistul nu are nimic de cîștigat dacă își distorsionează prognoza

Rezultate în afara eșantionului, 2020–2026



- Stînga: fiecare model cu cozi groase depășește scorul logaritmic al modelului Normal cu 0,13 pînă la 0,21 pe zi; diferențele dintre ele sînt mici
- Dreapta: ratele de depășire ale VaR 1%; ținta este linia punctată de la 1%

SFM_ch6_out_of_sample

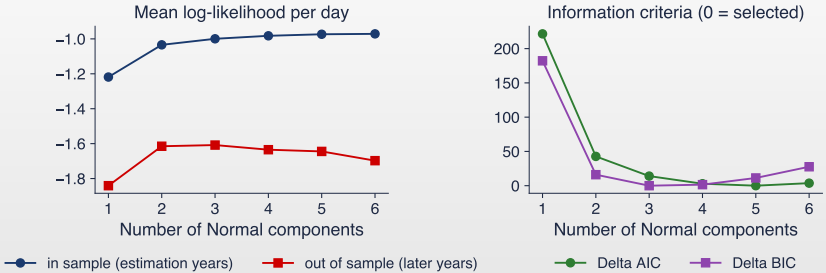
Interpretarea tabelului din afara eșantionului

	aștept.	Normală	t	skew-t	NIG	stabilă	ist.	cel mai bun AIC	cel mai bun scor
BET	16,8	17	5	5	5	3	5	NIG	Student-t
S&P 500	16,9	38	19	14	14	8	23	NIG	skewed-t
DAX	17,1	27	11	8	8	4	8	NIG	NIG
Bitcoin	24,5	23	7	7	7	1	10	GED	NIG

- Depășiri ale VaR 1% pe zilele de test; aștept.: numărul așteptat; ist.: simularea istorică (cuantila de 1% a eșantionului de estimare)
- Cîștigătorul AIC din 2000–2019 are cel mai bun scor logaritmic doar pentru DAX; în rest, cel mai bun scor aparține unui model apropiat în clasament
- BET: modelul Normal are 17 depășiri față de 16,8 așteptate, modelele cu cozi groase doar 5
  - perioada 2020–2026 a fost mai calmă pentru BET decît 2000–2019
- Numărul potrivit, dar din motive greșite: un model static nu poate urmări schimbările de volatilitate

SFM_ch6_out_of_sample

Supraajustare: amestecuri Normale cu 1 pînă la 6 componente



- S&P 500, estimare 2017–2018 ($n = 501$), test 2019–2026 ($n = 1\,939$); c componente au $3c - 1$ parametri
- În eșantion (albastru), calitatea ajustării crește cu fiecare componentă; în afara eșantionului (roșu) atinge maximul la $c = 3$ și apoi scade

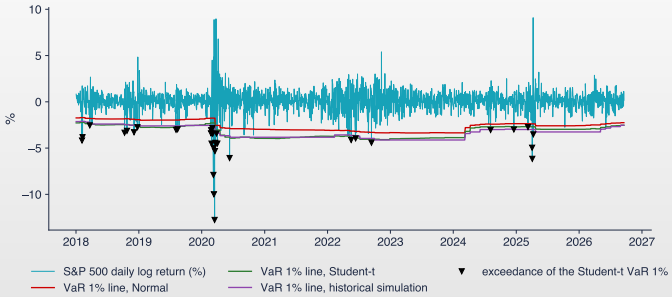
SFM_ch6_out_of_sample

Criteriile informaționale și supraajustarea

- Log-verosimilitatea medie pe zi, în eșantion → în afara eșantionului
  - $c = 1$ (Normală): $-1{,}219 \to -1{,}841$; $c = 3$: $-0{,}999 \to -1{,}608$; $c = 6$ (17 parametri): $-0{,}971 \to -1{,}698$
- BIC a ales $c = 3$, cel mai bun model în afara eșantionului; AIC a ales $c = 5$
  - cu $n = 501$, penalizarea AIC de 2 pe parametru este prea slabă pentru a opri componentele în plus
- Diferența dintre ajustarea în eșantion și cea din afara lui crește cu numărul de parametri: această diferență este supraajustarea
- Regula: eșantioane scurte, mulți parametri ⇒ preferați BIC sau AICc și păstrați întotdeauna o perioadă de test

SFM_ch6_out_of_sample

VaR 1% pe fereastră mobilă: reestimare lunară



- S&P 500: fiecare model este reestimat la fiecare 20 de zile de tranzacționare pe cele 1000 de zile anterioare; linia VaR 1% este desenată ca prag de pierdere $-\mathrm{VaR}$
- Graficul arată doar perioada 2018–2026; backtesting-ul de pe slide-ul următor acoperă 2003–2026. Triunghiurile negre: depășiri ale VaR Student-t; apar grupat (martie 2020)

SFM_ch6_out_of_sample

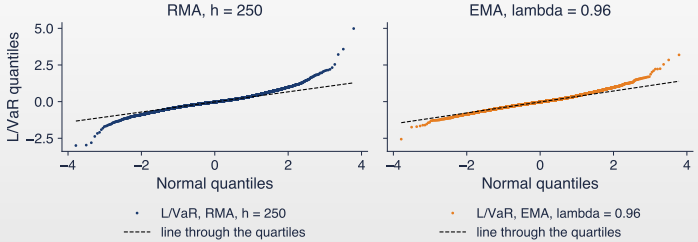
Backtesting pe fereastră mobilă: rezultatele

	Normală	Student-t	skewed-t	istorică
S&P 500, rata (%)	2,35	1,56	1,47	1,58
S&P 500, Kupiec p	$< 0{,}001$	$< 0{,}001$	$< 0{,}001$	$< 0{,}001$
S&P 500, maximul anual	36	31	28	29
BET, rata (%)	1,98	1,52	1,38	1,25
BET, Kupiec p	$< 0{,}001$	$< 0{,}001$	0,007	0,066

- Backtesting 2003–2026, $n = 5\,714$ zile pentru S&P 500; rata țintă 1%; cel mai slab an este 2008, pentru toate modelele și pentru ambele serii
- Cozile groase înjumătățesc excesul de depășiri, dar niciun model necondiționat nu ajunge la 1%
  - volatilitatea se schimbă mai repede decît se poate adapta o fereastră de 1000 de zile
- Remediul este un model pentru volatilitatea însăși: GARCH, Capitolele 8 și 9

SFM_ch6_out_of_sample

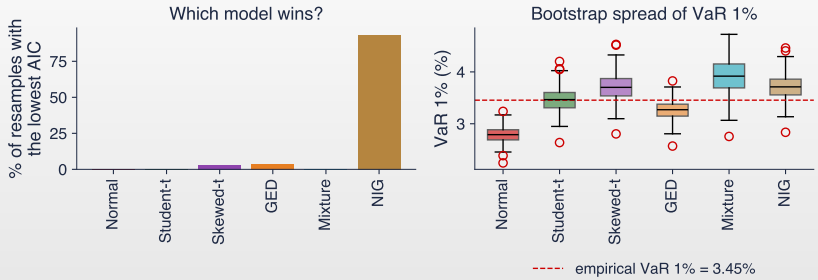
Graficul de fiabilitate VaR (SFEVaRqqplot)



- DAX din 2000: VaR 1% Normal $= -z_{\mathbf{0,01}}\hat{\sigma}_t = 2{,}326\,\hat{\sigma}_t$, cu $\hat{\sigma}_t$ din ultimele 250 de zile
  - RMA (rectangular moving average, medie mobilă simplă): $\hat{\sigma}_t^{\mathbf{2}} = \frac{\mathbf{1}}{\mathbf{250}} \sum_{j=\mathbf{1}}^{\mathbf{250}} r_{t-j}^{\mathbf{2}}$, ponderi egale
  - EMA (exponential moving average, medie mobilă exponențială): $\hat{\sigma}_t^{\mathbf{2}} = (1-\lambda) \sum_{j \geq \mathbf{0}} \lambda^j r_{t-\mathbf{1}-j}^{\mathbf{2}}$, $\lambda = 0{,}96$: fiecare zi mai veche are o pondere cu 4% mai mică
  - graficul QQ al lui L/VaR (pierderea împărțită la VaR) față de cuantilele Normale
- O dreaptă ar însemna un VaR fiabil; ambele se curbează în sus la capătul drept
  - pierderile dincolo de VaR sînt prea mari și prea frecvente (2,13% și 2,07% depășiri)

SFM_ch6_var_qqplot

Incertitudinea clasamentului modelelor



- S&P 500, 200 de reeșantionări bootstrap pe blocuri mobile (blocurile de 20 de zile păstrează volatility clustering); în fiecare reeșantionare sînt reestimate cele șase modele în afară de distribuția stabilă (Normală, Student-t, skewed-t, GED, amestec Normal, NIG)
- Stînga: NIG are cel mai mic AIC în 94% din reeșantionări; dreapta: dispersia bootstrap a VaR 1% pentru fiecare model

SFM_ch6_model_uncertainty

Două surse de incertitudine

- **Incertitudinea parametrilor**: același model, reestimat pe date reeșantionate
  - VaR 1% Student-t: intervalul bootstrap de 90% [3,14; 3,87]%; NIG: [3,34; 4,18]%
- **Incertitudinea de model**: modele diferite pe aceleași date
  - VaR 1% median: GED 3,27%, Student-t 3,47%, NIG 3,71%, amestec Normal 3,92%
  - dispersia dintre modele este comparabilă cu dispersia din interiorul unui model
- Chiar cînd cîștigătorul AIC este clar, cifra relevantă pentru risc rămîne incertă
- Măsurile de coadă construite pe baza entropiei oferă o altă cale de a rezuma această incertitudine: Pele, Lazar & Dufour (2017), Pele, Lazar & Mazurencu-Marinescu-Pele (2019)

Medierea modelelor și riscul de model

	date	Normală	interval (cozi groase)	mediat	interval (%)
BET	4,13	3,19	[3,74; 4,94]	4,11	32
S&P 500	3,45	2,80	[3,27; 4,54]	3,72	39
DAX	4,16	3,25	[3,77; 5,01]	4,24	33
Bitcoin	10,51	7,99	[9,85; 14,26]	9,85	45
TLV	5,10	4,01	[4,49; 5,41]	4,62	21
SNP	4,29	3,70	[4,16; 4,70]	4,25	13
BRD	4,50	3,81	[4,18; 4,74]	4,33	14

- VaR 1% în %; intervalul cozilor groase: cel mai mic și cel mai mare VaR 1% dintre cele șase modele nenormale, inclusiv amestecul Normal
  - intervalul (%): max/min -1, distanța relativă dintre modelele extreme
- Mediat: $\sum_i w_i \, \mathrm{VaR}_i$, cu ponderile Akaike w_i
  - un analog frecventist al medierii bayesiene a modelelor Hoeting et al. (1999)
- **Riscul de model**: amplitudinea rezultatelor între modelele plauzibile; autoritățile de supraveghere cer băncilor să îl măsoare Kerkhof, Melenberg & Schumacher (2010), Danielsson et al. (2016)

SFM_ch6_model_uncertainty

Etapele selecției modelului

- **1. Datele**: comparați cele mai mari randamente cu prețurile și cu știrile; eliminați doar erorile dovedite
- **2. Candidații**: alegeți-i după faptele stilizate (cozi, asimetrie), nu după comoditate
- **3. Ajustare și clasament**: ML, apoi AIC și BIC; teste LR pentru perechile imbricate, Vuong pentru celelalte
- **4. Verificare**: statistica AD și graficul QQ pentru primele modele din clasament; modelul trebuie să se potrivească bine în zona care vă interesează
- **5. Validare**: în afara eșantionului, cu funcția de pierdere potrivită scopului (depășiri pentru VaR)
- **6. Raportare**: cîștigătorul, modelele de pe locurile următoare și intervalul valorilor de risc între acestea

O întrebare de cercetare deschisă

- **Mai cîștigă NIG după ce eliminăm efectul de volatility clustering?**
  - o parte din cozile groase ale randamentelor zilnice provine din volatilitatea variabilă în timp (Capitolul 2)
  - randamentele împărțite la o prognoză de volatilitate (de exemplu EMA, $\lambda = 0{,}96$) au cozi mai subțiri; care candidat li se potrivește cel mai bine?
- Motivul pentru care este deschisă: răspunsul diferă de la o piață la alta, iar modelul de volatilitate este el însuși o alegere
- Punctul de plecare: pe randamentele brute, modelul cîștigător depinde de piață (harta ΔAIC din Secțiunea 4)
- Instrumentele AI pot accelera un astfel de studiu; nu înlocuiesc verificarea rezultatelor Wang et al. (2023)

Utilizări posibile ale instrumentelor AI

- **Literatura**: lista studiilor care compară distribuțiile randamentelor după filtrarea GARCH, cu eșantioanele și criteriile lor
- **Cod**: o primă versiune a buclei „standardizează, ajustează șapte modele, ordonează după AIC și BIC, testează în afara eșantionului”
- **Robustețe**: alte filtre de volatilitate, alte ferestre și alte piețe
- **Explicație**: o primă explicație a schimbării clasamentului după filtrare
- Exemplu de prompt
  - `Write Python code that divides daily log returns by an EMA volatility (lambda = 0.96), fits Normal, Student-t, skewed-t, GED and NIG by maximum likelihood, and reports AIC, BIC and Akaike weights.`

Verificări necesare

- Informația din viitor (look-ahead bias): volatilitatea zilei t trebuie să folosească doar randamentele pînă la $t-1$
- Numărul de parametri: k din AIC trebuie să includă fiecare parametru estimat, inclusiv pe cei ai filtrului de volatilitate, dacă au fost estimați
- Aceleași date pentru toate modelele: valorile AIC calculate pe eșantioane diferite nu se pot compara
- Parametrizările: scală sau abatere standard la Student-t; S0 sau S1 la distribuția stabilă
- Referințele: fiecare lucrare citată trebuie să existe; verificați DOI-ul

Idee de proiect

- **Întrebarea**: diferă cea mai bună distribuție pentru randamentele standardizate de cea pentru randamentele brute?
  - date: BET, S&P 500, DAX, Bitcoin și acțiunile BVB din datele cursului (EODHD)
- Pași
  - standardizați cu volatilitatea EMA; ajustați cei șapte candidați; tabel cu ΔAIC și ponderile Akaike
  - repetați backtesting-ul VaR 1% în afara eșantionului din acest capitol cu modelul standardizat
  - comparați grupurile de depășiri din 2008 și 2020 cu cele ale modelelor necondiționate
- Rezultat: un tabel, un grafic și un paragraf despre ce se schimbă și de ce
- Declarați orice folosire a AI și enumerați erorile AI pe care le-ați corectat

Idei de reținut

- Alegerea modelului începe cu date curate și un scop: corpul sau coada?
- ML estimează parametrii; testele LR compară modele imbricate (atenție la frontiere); Vuong le compară pe cele neimbricate
- AIC și BIC ordonează toți candidații; citiți Δ și ponderile Akaike
- Testele EDF: AD pune accentul pe cozi, KS nu; cu n mare toate modelele sînt respinse, deci ordonați modelele și examinați graficele
- Pe randamentele reale cîștigă NIG, GED și Student-t; distribuțiile Normală și stabilă pierd peste tot
- Validați în afara eșantionului și raportați intervalul VaR între modelele plauzibile

Formule de reținut

Mărimea	**Formula**
Log-verosimilitatea	$\ell(\theta) = \sum_t \ln f(r_t;\theta)$
Raportul de verosimilitate	$LR = 2(\ell_{\mathbf{1}} - \ell_{\mathbf{0}}) \approx \chi^{\mathbf{2}}(k_{\mathbf{1}} - k_{\mathbf{0}})$
AIC, BIC	$-2\ell(\hat\theta) + 2k, \quad -2\ell(\hat\theta) + k\ln n$
Ponderea Akaike	$w_i = e^{-\Delta_i/\mathbf{2}} / \sum_j e^{-\Delta_j/\mathbf{2}}$
Vuong	$V = \sqrt{n}\,\bar m / s_m, \quad m_t = \ln f_{\mathbf{1}}(r_t) - \ln f_{\mathbf{2}}(r_t)$
KS	$D_n = \sup_x \lvert F_n(x) - F(x)\rvert$
Anderson–Darling	$A^{\mathbf{2}} = n\int [F_n - F]^{\mathbf{2}}/[F(1-F)]\,dF$
Kupiec	$LR_{uc} = -2\ln \frac{(\mathbf{1}-\alpha)^{n-x}\alpha^x}{(\mathbf{1}-x/n)^{n-x}(x/n)^x}$
Pierderea pinball	$\frac{\mathbf{1}}{n}\sum_t(\alpha - 1\{r_t < q\})(r_t - q)$

Autoevaluare

- **Întrebare**: modelul A are $\ell = -1000$ cu $k = 3$, modelul B are $\ell = -998$ cu $k = 5$; care are AIC mai mic?
  - **Răspuns**: A: 2006; B: 2006; egalitate ($\Delta = 0$): cei doi parametri în plus aduc un cîștig egal exact cu penalizarea lor
- **Întrebare**: A este imbricat în B; este B semnificativ mai bun la 5%?
  - **Răspuns**: $LR = 2 \cdot 2 = 4 < \chi^2_{0,95}(2) = 5{,}99$: nu
- **Întrebare**: de ce nu poate un test KS cu parametri estimați să folosească tabelul KS clasic?
  - **Răspuns**: estimarea apropie F de date, deci D_n este prea mic; folosiți Lilliefors sau un bootstrap parametric
- Urmează: Capitolul 7, piețele eficiente, mersul aleator și testele variance ratio

Bibliografie (1/3)

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Amisano, G., Giacomini, R. (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics*, 25(2), 177–190.
- Anderson, T.W., Darling, D.A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Annals of Mathematical Statistics*, 23(2), 193–212.
- Anderson, T.W., Darling, D.A. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, 49(268), 765–769.
- Barndorff-Nielsen, O.E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, 24(1), 1–13.
- Borak, S., Härdle, W.K., López-Cabrera, B. (2013). *Statistics of Financial Markets: Exercises and Solutions*, 2nd ed. Springer.
- Box, G.E.P. (1976). Science and statistics. *Journal of the American Statistical Association*, 71(356), 791–799.
- Burnham, K.P., Anderson, D.R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd ed. Springer.
- Burnham, K.P., Anderson, D.R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33(2), 261–304.
- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223–236.
- Cramér, H. (1928). On the composition of elementary errors. *Scandinavian Actuarial Journal*, 1928(1), 13–74.
- Danielsson, J., James, K.R., Valenzuela, M., Zer, I. (2016). Model risk of risk models. *Journal of Financial Stability*, 23, 79–91.
- Diks, C., Panchenko, V., van Dijk, D. (2011). Likelihood-based scoring rules for comparing density forecasts in tails. *Journal of Econometrics*, 163(2), 215–230.

Bibliografie (2/3)

- **Eberlein, E., Keller, U. (1995).** Hyperbolic distributions in finance. ***Bernoulli*, 1(3), 281–299.**
- **Franke, J., Härdle, W.K., Hafner, C.M. (2019).** ***Statistics of Financial Markets: An Introduction*, 5th ed. Springer.**
- **Gneiting, T., Raftery, A.E. (2007).** Strictly proper scoring rules, prediction, and estimation. ***Journal of the American Statistical Association*, 102(477), 359–378.**
- **Hansen, B.E. (1994).** Autoregressive conditional density estimation. ***International Economic Review*, 35(3), 705–730.**
- **Hoeting, J.A., Madigan, D., Raftery, A.E., Volinsky, C.T. (1999).** Bayesian model averaging: a tutorial. ***Statistical Science*, 14(4), 382–417.**
- **Hurvich, C.M., Tsai, C.-L. (1989).** Regression and time series model selection in small samples. ***Biometrika*, 76(2), 297–307.**
- **Kerkhof, J., Melenberg, B., Schumacher, H. (2010).** Model risk and capital reserves. ***Journal of Banking & Finance*, 34(1), 267–279.**
- **Kon, S.J. (1984).** Models of stock returns – a comparison. ***Journal of Finance*, 39(1), 147–165.**
- **Kullback, S., Leibler, R.A. (1951).** On information and sufficiency. ***Annals of Mathematical Statistics*, 22(1), 79–86.**
- **Kupiec, P.H. (1995).** Techniques for verifying the accuracy of risk measurement models. ***Journal of Derivatives*, 3(2), 73–84.**
- **Lilliefors, H.W. (1967).** On the Kolmogorov–Smirnov test for normality with mean and variance unknown. ***Journal of the American Statistical Association*, 62(318), 399–402.**
- **Nelson, D.B. (1991).** Conditional heteroskedasticity in asset returns: a new approach. ***Econometrica*, 59(2), 347–370.**
- **Pearson, K. (1900).** On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. ***Philosophical Magazine*, 50(302), 157–175.**

Bibliografie (3/3)

- Pele, D.T., Lazar, E., Dufour, A. (2017). Information entropy and measures of market risk. *Entropy*, 19(5), 226.
- Pele, D.T., Lazar, E., Mazurencu-Marinescu-Pele, M. (2019). Modeling expected shortfall using tail entropy. *Entropy*, 21(12), 1204.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2), 461–464.
- Self, S.G., Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, 82(398), 605–610.
- Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics*, 19(2), 279–281.
- Stephens, M.A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69(347), 730–737.
- Stone, M. (1977). An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. *Journal of the Royal Statistical Society, Series B*, 39(1), 44–47.
- Stute, W., González Manteiga, W., Presedo Quindimil, M. (1993). Bootstrap based goodness-of-fit-tests. *Metrika*, 40(1), 243–256.
- Tsay, R.S. (2010). *Analysis of Financial Time Series*, 3rd ed. Wiley.
- Vuong, Q.H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57(2), 307–333.
- Wang, H., Fu, T., Du, Y., Gao, W., et al. (2023). Scientific discovery in the age of artificial intelligence. *Nature*, 620, 47–60.
- Wilks, S.S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Annals of Mathematical Statistics*, 9(1), 60–62.