









Modelling Financial Markets

Seminar 1: Stylised Facts and Returns — Hands-on



Daniel Traian PELE
Antoaneta Amza

Bucharest University of Economic Studies
IDA Institute Digital Assets
Blockchain Research Center
AI4EFin Artificial Intelligence for Energy Finance
Romanian Academy, Institute for Economic Forecasting
MSCA Digital Finance

What Can We Claim About Returns?

- Question of the seminar
  - **which stylised facts of returns survive inference that allows for heavy tails and volatility clustering?**
- Three types of exercises
  - **A**: calculation exercises, on paper
  - **B**: exercises on real data, solved in the notebook
  - **C**: open questions, e.g. correcting an answer given by an artificial intelligence (AI) tool
- Each exercise is marked
  - **[Solved]**: the full solution is presented in the seminar
  - **[Proposed]**: you solve it; the solution is discussed in the seminar

Which Exercises Come First?

- Today's exercises, in order
  - A1: aggregating returns over time and annualisation
  - A2(a): the depth of a drawdown (the fall from an earlier peak)
  - B1: descriptive statistics and tail thickness for three markets
  - A3: the Jarque–Bera and Ljung–Box tests, step by step
  - B2: heavy tails and the 1% Value-at-Risk (VaR) of BET-TR (BET Total Return, index), the Bucharest Stock Exchange BET (Bucharest Exchange Trading, index) index with dividends reinvested
  - A4: the kurtosis of a mixture of Normal distributions
  - B3: uncorrelated but not independent returns (Bitcoin)
  - B4: the Ljung–Box test robust to volatility clustering
  - B5: aggregating returns and approaching the Normal distribution
  - C1: correcting an AI answer with errors
- After lecture 1: tail estimation and robust autocorrelation bands
  - A5 and B6: the Hill estimator of the tail index
  - A6: the widening of the autocorrelation band under volatility clustering

What Do We Keep from Each Analysis?

- Outputs to produce
  - the answers to the A exercises
  - one figure and one table per B exercise
  - the corrected C1 code
  - one interpretation sentence per exercise
- Notebook for today: chapter1_seminar_notebook.ipynb (Colab); prices: data/market
  - each task names its notebook section
- Seminar: Prof. Daniel Traian Pele, danpele@ase.ro; Antoaneta Amza, PhD candidate, antoaneta.amza@csie.ase.ro

How Do We Read the Series and Measures?

- **JB**: Jarque–Bera (test).
- **ACF**: AutoCorrelation Function.
- **BVB**: Bursa de Valori București — the Bucharest Stock Exchange.
- **UTC**: Coordinated Universal Time (Temps Universel Coordonné).
- **BNR**: Banca Națională a României — the National Bank of Romania.
- **S&P 500**: Standard & Poor's 500.
- **OHLC**: Open, High, Low, Close.
- **US**: United States.
- **SPY**: Fund tracking the S&P 500 index.
- **ETF**: Exchange-Traded Fund.
- **VIX**: Cboe Volatility Index, the implied-volatility index of the S&P 500.

Exercise Map

Exercise	Question	Type	Model
A1	how do returns add up over time?	Solved	–
A2(a)	how deep is a drawdown?	Proposed	A1
B1	how heavy are the tails, and how fragile is the kurtosis?	Solved	–
A3	from a statistic to a p-value: JB and Ljung–Box	Solved	–
B2	how large is the VaR 1% of BET-TR?	Proposed	B1
A4	why does random volatility create heavy tails?	Solved	–
B3	are Bitcoin returns independent?	Solved	–
B4	which autocorrelations survive robust inference?	Solved	–
B5	do returns approach the Normal distribution with the horizon?	Proposed	B1, B3
C1	what is wrong in an AI answer?	Proposed	A3, B3, B4
After the lecture	A5 [Solved], B6 [Solved], A6 [Proposed] (tail heaviness, width of the ACF band)	review solved; solve A6	A4, B4
Extension packs	A1(c), A2(b–f), A7–A11, B1(d), B7–B15, C2, C3 (optional)	Mixed (marked)	named on each task

Data Used

Series	Price used	Period	Calendar	Returns
BET-TR	daily close	23 Sep 2014 – 18 Sep 2026	BVB trading days	2 992
Bitcoin	daily close (USD)	17 Sep 2014 – 18 Sep 2026	every day, 24:00 UTC	4 384
EUR/RON	BNR reference rate	4 Jul 2005 – 18 Sep 2026	Romanian working days	5 352
S&P 500	daily OHLC	2 Jan 1990 – 18 Sep 2026	US trading days	9 245
SPY ETF	OHLC, adjusted close (for dividends and splits)	2 Jan 2008 – 18 Sep 2026 (B15)	US trading days	–
VIX, 9-day VIX	daily close	C3 only	US trading days	–
Gold	spot, USD per ounce	2 Jan 1990 – 18 Sep 2026	working days	9 559

- Daily prices of the course data, one series per asset (field used: close / adjusted close)
- EUR/RON, the number of lei (RON) per euro (EUR), is the official BNR reference rate, not a series of market quotes
- Units: returns in decimals in the code, reported in % on the slides
- Calendars: each series on its own calendar
  - multi-series analyses first align the *prices* on common days

Files for SPY and the VIX Indices

- CSV (Comma-Separated Values) files supplied in `data/market`; all end on 18 Sep 2026.
- Counts below are price observations, before differencing.

File	First date	Rows	Exercise
SPY.US.csv	2008-01-02	4 708	B15, C3
VIX.INDX.csv	1990-01-02	9 284	C3
VIX9D.INDX.csv	2011-01-03	3 951	C3

- Fields: `open`, `high`, `low`, `close`, `adjusted_close`, `volume`, indexed by `date`.
- Use `read_market("SPY.US")` from `mfm_data`; select SPY dates from 2008. Use `"VIX.INDX"` and `"VIX9D.INDX"` for the volatility indices.

BVB stocks for C2

- Supplied files: `data/market/<ticker>.RO.csv`; last date: 18 Sep 2026.
- C2 uses `volume` and raw `close` in lei (RON). Counts are price rows before filtering.

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

Setup: Run and Verify

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

- Expected output: first return date and count

Series	first return	T
S&P 500	1990-01-03	9 245
BET-TR	2014-09-24	2 992
Bitcoin	2014-09-18	4 384
EUR/RON	2005-07-05	5 352

- Conventions in the code
  - moments: `scipy.stats.skew` and `kurtosis` (excess kurtosis, no small-sample correction)
  - standard deviations with `ddof=1`
  - quantiles with `numpy.quantile` (linear)
- If a count differs, stop and check the series, the price type and the start date

Prices and Returns

- Indices and prices
  - $t = 0, 1, \ldots, T$: the trading-day index; $t - 1$ is the previous day
  - P_t: the closing price on day t; T: the number of returns
- **Simple return**: the relative price change over one day

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

- **Log return**: the natural logarithm (ln) of the price ratio

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

  - $r_t \approx R_t$ for small moves; $r_t < R_t$ whenever $R_t \neq 0$
  - one worked line: $P_0 = 100$, $P_1 = 102$: $R_1 = 2.00\%$, $r_1 = \ln 1.02 = 1.98\%$

Mean, Standard Deviation, Standardisation

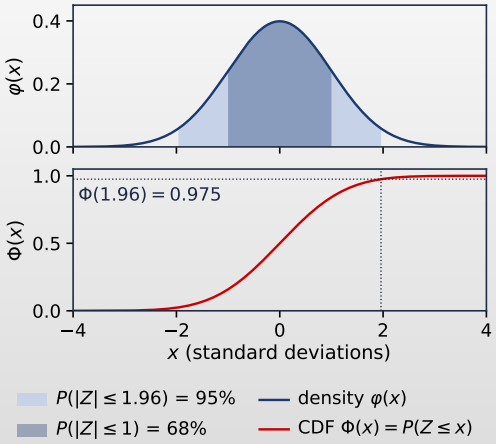
- Writing conventions
  - the hat: $\hat{\theta}$ is the sample estimate of a population quantity θ
  - $\sum_t$: the sum over all days $t = 1, \ldots, T$
- **Sample mean** $\bar{r} = \frac{1}{T} \sum_t r_t$: the average daily return
- **Sample standard deviation** (daily volatility): the typical spread of returns around the mean

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

- Demeaned return $x_t = r_t - \bar{r}$: the deviation from the mean
- Standardised return $z_t = (r_t - \bar{r})/\hat{\sigma}$: the number of standard deviations from the mean
  - $z_t = -3$: a day three standard deviations below the mean
- **i.i.d.**: independent and identically distributed, the reference assumption of the classical tests
  - today's return does not depend on past returns, and its distribution is the same every day

The Normal Curve

- $Z \sim N(0,1)$
  - $\varphi(x)$: the height of the curve at x
  - $\Phi(x)$: the area under the curve to the left of x
- Shaded areas in the chart
  - $P(|Z| \leq 1) = 68\%$
  - $P(|Z| \leq 1.96) = 95\%$
- Φ^{-1}: the inverse of Φ, i.e. the $N(0,1)$ quantiles
  - $\Phi^{-1}(0.975) = 1.96$, so Z lies in $[-1.96, 1.96]$ with probability 95%
- Standardisation: if $X \sim N(\mu, \sigma^2)$, then $Z = (X - \mu)/\sigma \sim N(0,1)$



Aggregation over Time and Annualisation

- **Aggregation over time**: the return over days $1, \ldots, n$, from P_0 to P_n

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

  - $\prod$: the product of the factors; simple returns compound (the factors $1 + R_t$ multiply)
  - log returns add up: $\sum_{t=1}^{n}(\ln P_t - \ln P_{t-1}) = \ln P_n - \ln P_0$, because the intermediate terms cancel ($+\ln P_1$ with $-\ln P_1$, etc.)
  - across assets, in a portfolio, it is the simple returns that add up with weights (A1(c))
- **Annualisation**: A is the number of return intervals in a year

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

  - $A = 252$ (BET-TR, S&P 500, EUR/RON: trading days), $A = 365$ (Bitcoin, traded every day)
  - $A\sigma^2$ assumes uncorrelated returns with constant variance (A7: the autocorrelated case)
  - example: $\hat{\sigma} = 1\%$ a day and $A = 252$: $\hat{\sigma}_{\text{ann}} = 15.87 \times 1\% = 15.9\%$

Skewness, Kurtosis, Autocorrelation

- **Skewness** $\hat{S} = \frac{1}{T}\sum_t z_t^3$: the mean cube of the standardised returns
  - $\hat{S} = 0$ for a symmetric distribution
  - $\hat{S} < 0$: extreme losses are larger, in absolute value, than extreme gains
- **Kurtosis** $\hat{K} = \frac{1}{T}\sum_t z_t^4$: the mean fourth power, driven mostly by the extreme days
  - $\hat{K} = 3$ for the Normal distribution
  - **excess kurtosis** $\hat{K} - 3$: 0 for the Normal distribution, positive for heavy tails
- **Autocorrelation** at lag k: the correlation between returns k days apart

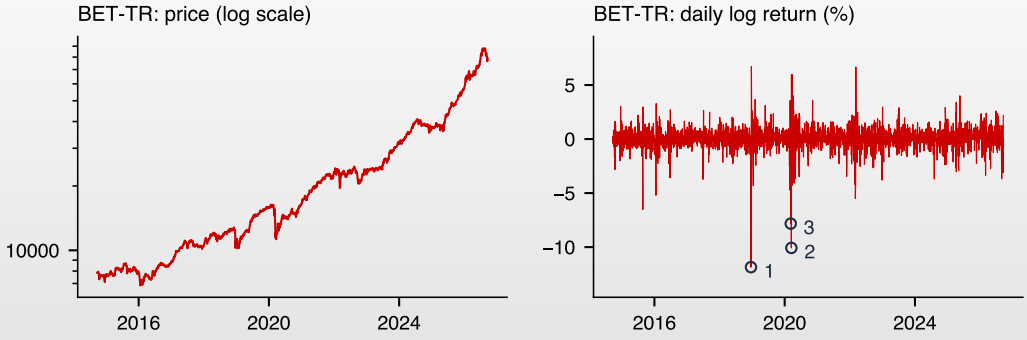
$$\hat{\rho}_k = \frac{\sum_{t=k+1}^{T} x_t x_{t-k}}{\sum_{t=1}^{T} x_t^2}$$

  - $k = 1, 2, \ldots$: the lag, in days; $\hat{\rho}_k \in [-1, 1]$; ρ_k without a hat: the population value
  - $\hat{\rho}_k > 0$: a day above the mean tends to be followed, k days later, by a day above the mean
  - $\hat{\rho}_k \approx 0$: no linear link between returns k days apart

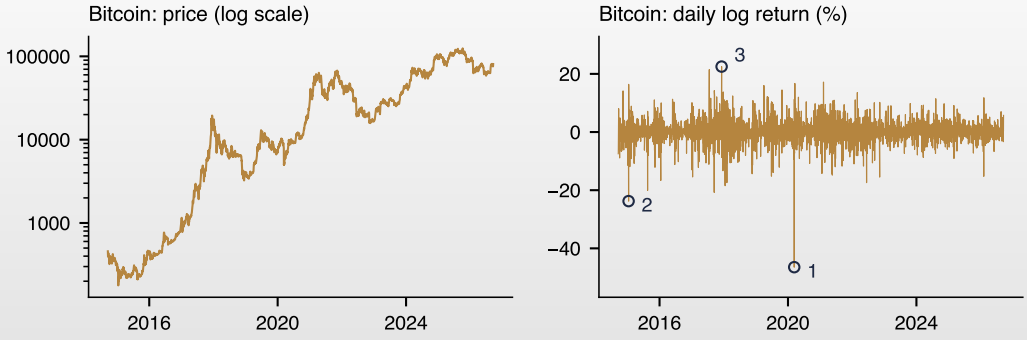
Statistical Tests

- **The χ^2_m distribution** (chi-squared with m degrees of freedom)
  - the distribution of the sum of squares of m independent $N(0,1)$ variables; its mean is m
  - 5% critical values: 3.84 ($m = 1$), 5.99 ($m = 2$), 11.07 ($m = 5$), 18.3 ($m = 10$)
- **Critical value** at 5%: the value exceeded with probability 5% under H_0
  - reject H_0 at the 5% level if the statistic exceeds the critical value
- **p-value**: probability, under H_0, of a statistic at least as extreme as the one observed
  - $p < 0.05$ is the same as exceeding the 5% critical value
  - it is not the probability that H_0 is true

BET-TR: Prices, Returns and Extreme Days



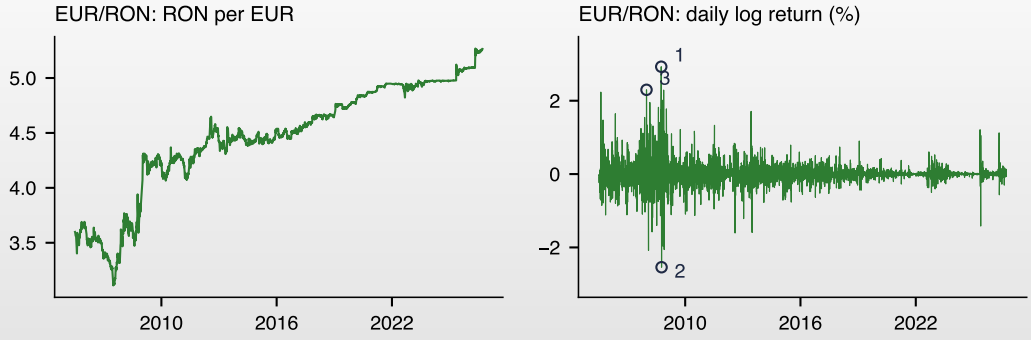
Bitcoin: Prices, Returns and Extreme Days



The three days with the largest absolute return

1: 12 Mar 2020, -46.5% 2: 14 Jan 2015, -23.8% 3: 07 Dec 2017, +22.5%

EUR/RON: Prices, Returns and Extreme Days



The three days with the largest absolute return

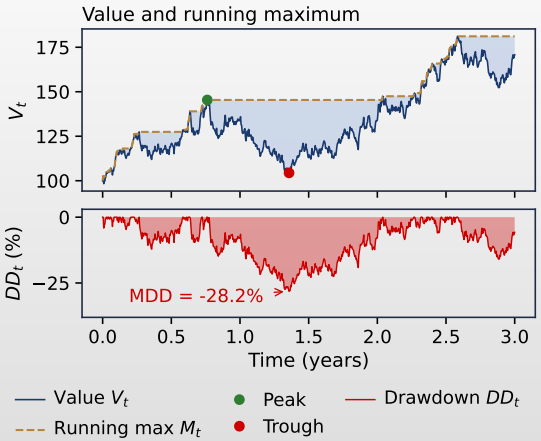
1: 03 Oct 2008, +2.9% 2: 09 Oct 2008, -2.5% 3: 28 Dec 2007, +2.3%

Stylised Facts and Market Episodes

- **Stylised fact**: a statistical property shared by many assets, markets and periods Cont (2001)
- Today's four stylised facts
  - **heavy tails**: extreme days are more frequent than under the Normal distribution with the same variance
  - **volatility clustering**: large moves follow large moves, of either sign
  - **no linear autocorrelation**: today's return is not a linear function of past returns, $\hat{\rho}_k \approx 0$
  - **aggregational Gaussianity**: returns over longer horizons have tails closer to those of the Normal distribution
- Every extreme day is a dated event
  - the bank-asset tax (19 Dec 2018)
  - the crash of March 2020, at the start of the COVID-19 (coronavirus disease 2019) pandemic
  - the 2008 depreciation of the leu
- Calm and turbulent periods alternate in all three series
- EUR/RON is almost flat after 2013

MFM_ch1_seminar_tests

The Drawdown



- **Running peak**: the highest value reached up to time t

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

  - V_t: the portfolio value or price at time t
- **Drawdown**: the relative loss from the running peak

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

  - $DD_t = 0$ at every new high; the plot of DD_t is the underwater curve
- **Maximum drawdown** (MDD) $MDD = \min_t DD_t$
  - the largest fall from a peak to a later trough
  - in the chart: -28.2%, from the green dot to the red dot
- **CAGR** (compound annual growth rate)
  $= (V_n/V_0)^{1/Y} - 1$: the constant annual growth rate
  - Y: the number of years between V_0 and V_n

The Sharpe, Sortino and Calmar Ratios

- **Sharpe ratio**: the annual excess return per unit of annual volatility

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

  - $\bar{R}$, $\hat{\sigma}_R$: the mean and standard deviation of simple returns; r_f: the annual risk-free rate
  - $SR = 0.5$: each percentage point of annual volatility earns 0.5 percentage points of excess return

- **Sortino ratio** Sortino & van der Meer (1991): penalises only the losing days

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

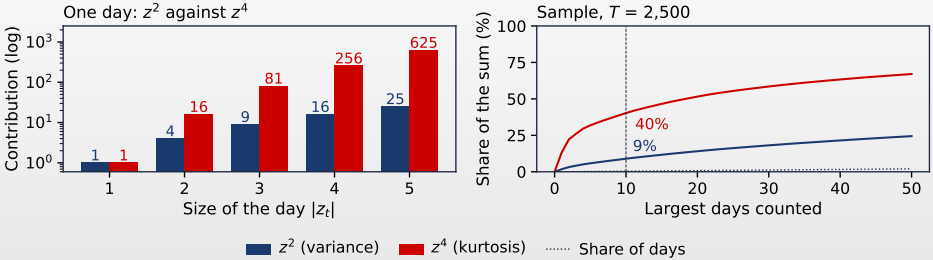
  - $\hat{\sigma}_D$: the downside semi-deviation; days with $R_t \geq 0$ contribute 0
  - Sortino $>$ Sharpe alone does not establish asymmetry: for a symmetric, near-zero-mean return, downside deviation is about $\sigma/\sqrt{2}$, so Sortino is already about 1.4 times Sharpe

- **Calmar ratio** $=$ CAGR$/|MDD|$ (previous slide)
  - annual growth per unit of maximum loss from a peak
  - it depends on a single episode, so it is descriptive, with no simple standard error

Moments of Order k and the Kurtosis

- **Moments** of a random variable r with mean μ and standard deviation σ
  - moment of order k: $\mathbb{E}[r^k]$; central moment of order k: $\mathbb{E}[(r-\mu)^k]$
  - standardised moment of order k: $\mathbb{E}[z^k]$, with $z = (r-\mu)/\sigma$
  - $\mu = \mathbb{E}[r]$ is the moment of order 1; the variance $\sigma^2 = \mathbb{E}[(r-\mu)^2]$ is the central moment of order 2
  - the skewness $S = \mathbb{E}[z^3]$ and the kurtosis $K = \mathbb{E}[z^4]$ are the standardised moments of order 3 and 4
  - **finite fourth moment**: $\mathbb{E}[r^4] < \infty$; otherwise $\hat{K}$ does not settle down as T grows
- **Sample kurtosis** $\hat{K} = \frac{1}{T}\sum_t z_t^4$ and the excess $\hat{K} - 3$
- Why the fourth moment measures the tails
  - a day's contribution grows very fast with $|z_t|$: $1^4 = 1$, $2^4 = 16$, $3^4 = 81$, $10^4 = 10{,}000$
  - so $\hat{K}$ is driven by the extreme days, not by ordinary days
  - one day with $z_t = 10$ alone adds $10^4/T$ to $\hat{K}$: $+4$ for $T = 2{,}500$

How Extreme Days Drive the Kurtosis



- Left: the contribution of a single day to $\sum_t z_t^2$ (variance) and to $\sum_t z_t^4$ (kurtosis), on a log scale
  - a day with $|z_t| = 5$ counts 25 times as much as a day with $|z_t| = 1$ in the variance, but 625 times as much in the kurtosis
- Right: a simulated heavy-tailed sample, $T = 2{,}500$; the share of the largest days in the two sums
  - the 10 largest days (0.4% of days) give 9% of $\sum_t z_t^2$ but 40% of $\sum_t z_t^4$
  - this is why $\hat{K}$ moves a lot when a few days are removed (B1)
- **What do you think**: which statistic changes more if we remove the most extreme day of the sample, the standard deviation or the kurtosis?
  - **Answer**: the kurtosis, because the day contributes z_t^4, not z_t^2

Quantiles and a Robust Tail Measure

- **Quantile** of order p, written q_p: the value below which a fraction p of the returns lies

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

  - F: the distribution function; empirically, q_p is the value at position pT in the returns sorted in increasing order
  - $q_{0.5}$ is the median; $q_{0.25}$ and $q_{0.75}$ are the quartiles
- Robust alternative Kim & White (2004): the Crow–Siddiqui quantile excess kurtosis, denoted K_{CS}

$$K_{CS} = \frac{q_{0.975} - q_{0.025}}{q_{0.75} - q_{0.25}} - 2.91$$

  - numerator: the width of the range holding 95% of the days; denominator: the width of the middle 50%
  - 2.91 is the value of the ratio for the Normal distribution, so the measure is 0 for the Normal distribution
  - positive value: the tails are long relative to the centre, longer than for the Normal distribution
  - it depends only on quantiles, so a few extreme days barely move it

The Jarque–Bera Test

- **Jarque–Bera** Jarque & Bera (1980): measures how far the skewness and the excess kurtosis are from their Normal values, both 0

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

  - 6 and 24: the variances of $\sqrt{T}\,\hat{S}$ and $\sqrt{T}(\hat{K}-3)$ under H_0, for large T
  - each term is therefore the square of an approximately $N(0,1)$ variable
  - $JB \geq 0$, and $JB = 0$ only if $\hat{S} = 0$ and $\hat{K} = 3$
- H_0: i.i.d. returns with the Normal distribution
  - under H_0, $JB \sim \chi^2_2$ for large T (two terms); 5% critical value 5.99
  - a large JB: the skewness or the tails do not fit the Normal distribution
- Example: $T = 1\,000$, $\hat{S} = 0$, $\hat{K} = 6$: $JB = 1\,000 \times 9/24 = 375 > 5.99$, reject H_0

Value-at-Risk and the Normal Distribution

- **Value-at-Risk (VaR) 1%** over one day: the loss threshold exceeded with probability 1%

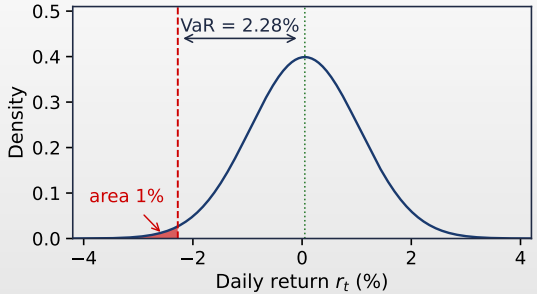
$$P\left(L_t > \mathrm{VaR}_\alpha\right) = \alpha, \qquad L_t = -r_t, \qquad \mathrm{VaR}_\alpha = -q_\alpha$$

  - L_t: the loss on day t; α: the VaR level, here $1\% = 0.01$
  - q_α: the α-quantile of the return; the minus sign turns the negative quantile into a positive loss
  - VaR is not the maximum loss: on about 1% of days the loss exceeds it, and VaR does not say by how much
- If the return has the Normal distribution $N(\mu, \sigma^2)$

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

  - $\Phi^{-1}(0.01) = -2.326$: 1% of the values of an $N(0,1)$ variable lie below -2.326
  - example: $\mu = 0.05\%$, $\sigma = 1\%$: $q_{0.01} = 0.05 - 2.326 = -2.28\%$, so VaR $1\% = 2.28\%$
- **Empirical VaR**: minus the sample 1% quantile
  - on the sample that produced it, it is breached on 1% of days by construction

The Quantile and VaR on a Density



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

- **Density** f: the area under the curve between a and b is $P(a < r_t \leq b)$; the total area is 1
- **Quantile** $q_{0.01}$: the point that leaves an area of 1% to its left (the red area)

$$\int_{-\infty}^{q_{0.01}} f(x)\,dx = 0.01$$

- **VaR 1%** $= -q_{0.01}$: the distance from 0 to the quantile, expressed as a positive loss
  - with heavier tails the 1% area sits further left and VaR grows (B2)

The Student-t Distribution

- **Student-t** with $\nu > 0$ degrees of freedom: a symmetric distribution with heavier tails than the Normal distribution
  - $P(|X| > x)$ decays like $x^{-\nu}$ as x grows: the smaller ν, the heavier the tails
  - as $\nu \to \infty$ it becomes the Normal distribution
- Moments: $\mathbb{E}|X|^p$ is finite only for $p < \nu$
  - the variance exists for $\nu > 2$ and equals $\nu/(\nu - 2)$; the kurtosis exists for $\nu > 4$
- Quantiles: $t_\nu^{-1}(\alpha)$ is the α-quantile of the standard Student-t
  - $t_3^{-1}(0.01) = -4.54$
  - rescaling to unit variance: divide by the standard deviation of the distribution, $\sqrt{\nu/(\nu-2)}$
  - for $\nu = 3$: $\sqrt{3/(3-2)} = \sqrt{3}$, so the standardised quantile is $-4.54/\sqrt{3} = -2.62$
  - it is deeper than -2.33, the 1% quantile of the Normal distribution
- Fitting: by maximum likelihood, for the variable $m + s\,Y$ with Y a standard Student-t
  - m: the location (centre); $s > 0$: the scale (width); `scipy.stats.t.fit`

The Student-t Density and the Tail Index

- **Standard Student-t density** with ν degrees of freedom

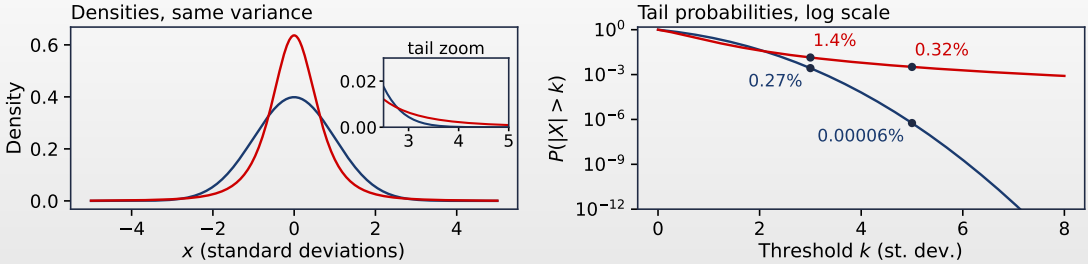
$$f_\nu(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

  - Γ: the gamma function, $\Gamma(n) = (n-1)!$ for integer n; the leading fraction makes the total area 1
  - for large $|x|$, $f_\nu(x)$ decays like $|x|^{-(\nu+1)}$, a power of $|x|$
  - the Normal density decays like $e^{-x^2/2}$, much faster: hence the difference in the tails
- **Tail index** α: how fast the probability of a large move falls

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

  - $\propto$: proportional to; a small α means heavy tails; for the Student-t, $\alpha = \nu$; not the VaR level α
  - $\mathbb{E}|r_t|^p$ is finite only for $p < \alpha$: with $\alpha \leq 4$ the kurtosis does not exist
- After the lecture
  - A5 and B6 estimate α with the Hill estimator; A6 derives the variance of $\hat{\rho}_k$

The Normal and the Student-t with the Same Variance



- Left: $N(0,1)$ and the Student-t with $\nu = 3$, rescaled to unit variance (divided by $\sqrt{3}$)
  - same variance, but the Student-t has a higher peak, thinner shoulders and heavier tails (zoom)
- Right: $P(|X| > k)$ on a log scale
  - $k = 3$: 1.4% for the Student-t against 0.27% for the Normal distribution, 5 times as likely
  - $k = 5$: 0.32% against 0.00006%, more than 5,000 times as likely

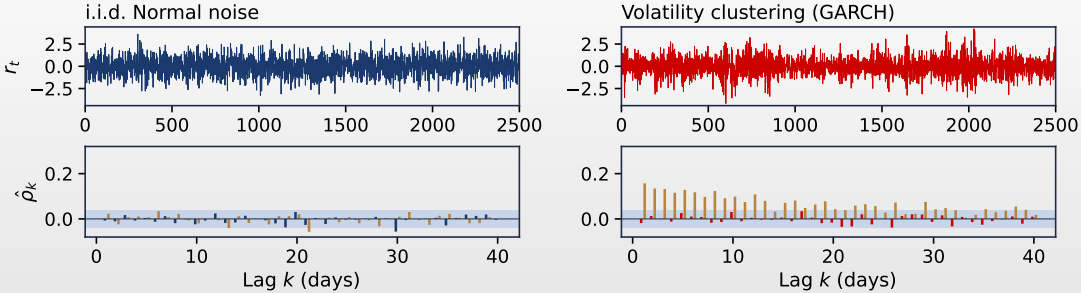
The ACF and the Ljung–Box Test

- **ACF** (autocorrelation function): the sequence $\hat{\rho}_k$, $k = 1, 2, \ldots$, plotted against the lag
  - for i.i.d. returns $\hat{\rho}_k \approx N(0, 1/T)$, so the standard deviation of $\hat{\rho}_k$ is $1/\sqrt{T}$
  - 95% band: $\pm 1.96/\sqrt{T}$ ($T = 2\,500$: ± 0.039); a $\hat{\rho}_k$ outside it is significant at 5%
- **Volatility clustering** in the ACF: $|r_t|$ and r_t^2 are autocorrelated although r_t is not
- **Ljung–Box** Ljung & Box (1978): tests the first m autocorrelations jointly

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

  - m: the number of lags tested; H_0: $\rho_1 = \cdots = \rho_m = 0$
  - for i.i.d. returns $Q(m) \sim \chi^2_m$; a Q above the critical value: at least one autocorrelation is non-zero
  - on r_t it tests linear predictability; on r_t^2 it tests volatility clustering

The ACF of i.i.d. Noise and under Volatility Clustering



- The series on the right, simulated from a GARCH model (generalised autoregressive conditional heteroskedasticity): $r_t = \sigma_t \varepsilon_t$, $\sigma_t^2 = 0.05 + 0.10\, r_{t-1}^2 + 0.85\, \sigma_{t-1}^2$, ε_t i.i.d. $N(0,1)$
  - a day with a large $|r_{t-1}|$ raises today's volatility: calm and turbulent periods alternate
- The ACF of r_t stays inside the band $\pm 1.96/\sqrt{T} = \pm 0.039$ ($T = 2{,}500$) for both series
- The ACF of $|r_t|$: inside the band for i.i.d. noise, far above it for the series on the right (0.16 at lag 1)

The ARCH-LM Test

- **ARCH-LM** (ARCH Lagrange Multiplier, test) Engle (1982): the Lagrange multiplier (LM) test for ARCH (autoregressive conditional heteroskedasticity) effects; tests whether today's squared return depends on past squared returns, i.e. volatility clustering
- The auxiliary regression, estimated by ordinary least squares (OLS)

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

  - q: the number of lags; $a_0, \ldots, a_q$: the coefficients; u_t: the error
  - H_0: $a_1 = \cdots = a_q = 0$, i.e. no volatility clustering
- The statistic $LM = (T - q)R^2$, compared with χ_q^2
  - $R^2 \in [0, 1]$: the share of the variation of x_t^2 explained by past squares
  - $T - q$: the number of observations used in the regression
  - example: $T = 1\,000$, $q = 5$, $R^2 = 0.02$: $LM = 995 \times 0.02 = 19.9 > 11.07$, the 5% critical value of χ_5^2: clustering

The Mixture of Normal Distributions

- **Mixture of Normal distributions** Clark (1973): the return is a Normal shock scaled by a random volatility

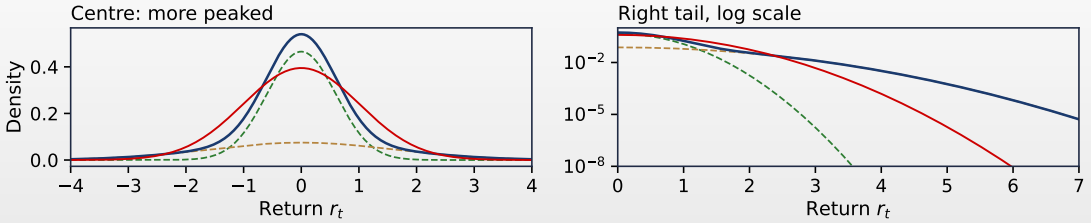
$$r_t = \sigma_t \varepsilon_t$$

  - ε_t: the shock, i.i.d. $N(0,1)$ and independent of σ_t
  - $\sigma_t > 0$: the volatility of day t, itself random (for example, driven by the pace of information arrival)
- The population kurtosis of r_t

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

  - 3 is $\mathbb{E}\varepsilon^4$, the kurtosis of $N(0,1)$
  - $K = 3$ only if σ_t is constant; the more σ_t varies, the larger K
  - heavy tails from Normal shocks (A4)

A Mixture of Two Normal Distributions



- Calm days, with probability 0.7: $\sigma_t = 0.6$; turbulent days, with probability 0.3: $\sigma_t = 1.6$

$$f(x) = 0.7\,\varphi_{\mathbf{0.6}}(x) + 0.3\,\varphi_{\mathbf{1.6}}(x), \qquad \varphi_s(x) = \frac{1}{s}\,\varphi\Big(\frac{x}{s}\Big)$$

- φ_s: the density of $N(0, s^{\mathbf{2}})$; mixture variance $0.7 \times 0.6^{\mathbf{2}} + 0.3 \times 1.6^{\mathbf{2}} = 1.02$; $K = 3\,\mathbb{E}\sigma^{\mathbf{4}}/(\mathbb{E}\sigma^{\mathbf{2}})^{\mathbf{2}} = 5.9$
- more very small and more very large days than under the Normal distribution with the same variance
- $P(|r_t| > 4)$: 0.37% in the mixture against 0.0075% for the Normal distribution

Inference under Volatility Clustering

- Under volatility clustering, $\mathrm{Var}(\hat{\rho}_k) > 1/T$
  - the band $\pm 1.96/\sqrt{T}$ is too narrow and $Q(m)$ rejects H_0 too often
- The variance of $\hat{\rho}_k$ is estimated directly from the data Lobato, Nankervis & Savin (2001)

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

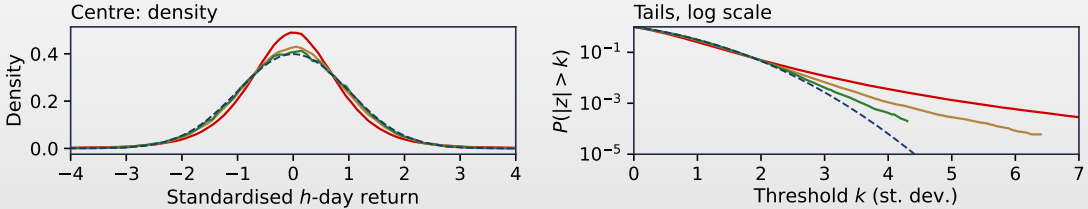
  - for i.i.d. returns $\hat{\tau}_k \approx 1/T$; $T\hat{\tau}_k$ says how many times the variance of $\hat{\rho}_k$ exceeds $1/T$
  - robust band $\pm 1.96\sqrt{\hat{\tau}_k}$
- Robust Ljung–Box (B4): each $\hat{\rho}_k^2$ is divided by its own estimated variance

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

  - the same H_0 as $Q(m)$, still compared with χ^2_m
  - example: $T\hat{\tau}_1 = 4$ makes the robust band $\sqrt{4} = 2$ times wider than $\pm 1.96/\sqrt{T}$

Aggregational Gaussianity

- **The h-day log return**: the sum of h consecutive daily returns, $r_t^{(h)} = r_t + r_{t-\mathbf{1}} + \cdots + r_{t-h+\mathbf{1}}$
  - h: the horizon, in days; κ: the excess kurtosis of the daily return
  - for i.i.d. returns the h-day excess kurtosis is $\kappa/h \to 0$: the sum approaches the Normal distribution



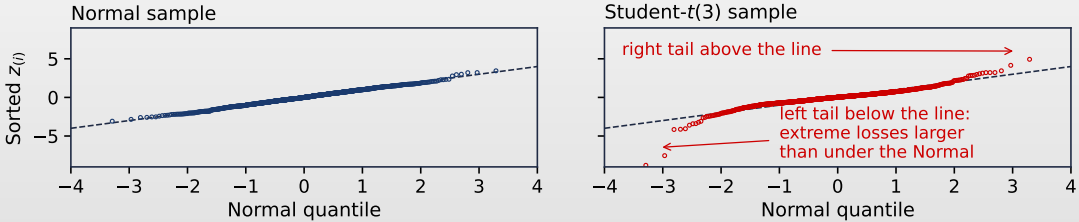
- Simulated i.i.d. heavy-tailed daily returns ($\kappa = 6$), summed over $h = 1, 5, 20$ days and standardised
  - the excess kurtosis falls from 6 to 1.2 and to 0.3; the tail (right) approaches that of the Normal distribution
  - under volatility clustering the approach is slower than κ/h (B5)

The QQ Plot

- **QQ (quantile–quantile) plot**: the data quantiles against the quantiles of $N(0,1)$, one point per day

$$\left(\Phi^{-1}\left(\frac{i-0.5}{T}\right),\ z_{(i)}\right), \qquad i = 1, \ldots, T$$

  - $z_{(1)} \leq \cdots \leq z_{(T)}$: the standardised returns sorted in increasing order; $(i-0.5)/T$: the matching proportion



- On the 45-degree line: the Normal distribution (left); an S shape: heavy tails (right, Student-t with $\nu = 3$)
  - the left tail lies below the line and the right tail above it: the extreme data quantiles exceed the Normal quantiles

A1 [Solved]: Returns and Annualisation — Task

- **Questions**
  - which return adds up over time?
  - how do daily numbers become annual ones?
- **Inputs**
  - a share closes at 100, 110, 99, 104.5 on days 0–3 and pays no dividend
  - an index has a daily mean log return of 0.04% and a daily volatility of 1.2%
  - Bitcoin has a daily volatility of 3.5%
  - returns are serially uncorrelated with constant variance
- **Tasks**
  1. Compute the three simple returns R_t and the three log returns r_t, all in %.
  2. Compute the 3-day returns $P_3/P_0 - 1$ and $\ln(P_3/P_0)$ and compare them with $\sum_t R_t$ and $\sum_t r_t$.
  3. Annualise the index mean log return and volatility with $A = 252$.
  4. Annualise the Bitcoin volatility and justify the choice of A.
- **Report**: a table day, price, R_t, r_t and the three annual numbers
- Notebook: section A1

Back Exercise list

A1 [Solved]: Solution (1/2)



A1 [Solved]: Solution (2/2)

t	P_t	R_t	r_t
0	100	–	–
1	110	+10.00%	+9.53%
2	99	−10.00%	−10.54%
3	104.5	+5.56%	+5.41%
sum		5.56%	4.40%
$0 \to 3$		4.50%	4.40%

- The sum of the log returns is $\ln(104.5/100)$: the intermediate terms $\ln P_t$ cancel
- Simple returns compound: $1.1 \times 0.9 \times 1.0556 - 1 = 4.5\%$

- Index
  - annualised mean log return $252 \times 0.04\% = 10.08\%$
  - volatility $\sqrt{252} \times 1.2\% = 15.87 \times 1.2\% = 19.05\%$
- Bitcoin trades every day: $\sqrt{365} \times 3.5\% = 66.9\%$
  - with $\sqrt{252}$ one reports 55.6%, 17% too low

MFM_ch1_seminar_tests

A2(a) [Proposed]: The Depth of a Drawdown — Task

- **Model**: A1 [Solved]; notebook section A2
- **Questions**
  - what is the worst loss from a previous peak?
  - how does it compare with the growth rate?
- **Inputs**
  - year-end portfolio values $V_0, \ldots, V_5 = 100, 120, 90, 110, 130, 117$
  - 5-year CAGR $(V_5/V_0)^{1/5} - 1 = 3.19\%$: the constant annual rate that takes V_0 to V_5
- **Tasks**: complete the worksheet
  1. Compute the running peak $M_t = \max_{s \le t} V_s$, the largest V_s over years $s = 0, \ldots, t$.
  2. Compute the drawdown $DD_t = V_t/M_t - 1 \le 0$, the relative loss from the peak.
  3. Find the maximum drawdown $MDD = \min_t DD_t$ and the year in which it occurs.
  4. Compute the Calmar ratio CAGR$/|MDD|$, the annual growth per unit of maximum loss.

Year	0	1	2	3	4	5
V_t	100	120	90	110	130	117
M_t	100	120	...	...	...	...
DD_t	0%	0%	...	...	...	...

- **Interpretation**: why is the fall to 117 in year 5 not the maximum drawdown?

Back Exercise list

B1 [Solved]: Descriptive Statistics — Task

- **Questions**
  - how heavy are the tails of BET-TR, Bitcoin and EUR/RON?
  - how much does the kurtosis depend on a few days?
- **Data**: `r_bet`, `r_btc`, `r_eur` from the setup
  - $T = 2\,992,\ 4\,384,\ 5\,352$; $A = 252,\ 365,\ 252$
- **Tasks**
  1. Compute the mean, the standard deviation, the annualised volatility, the skewness, the excess kurtosis, the minimum and the maximum of each series.
  2. Compute the JB statistic (definition) and its two terms.
  3. Compare JB with 5.99, the 5% critical value of χ^2_2.
  4. Recompute each excess kurtosis after removing the day with the largest $|r_t|$, and after removing the five largest.
  5. Recompute the Crow–Siddiqui quantile excess kurtosis K_{CS} (definition) on the same trimmed samples.
- **Report**: one table for steps 1–3 and a chart of kurtosis against the number of days removed
- **Interpretation**: which measure of tail heaviness can you report with confidence?

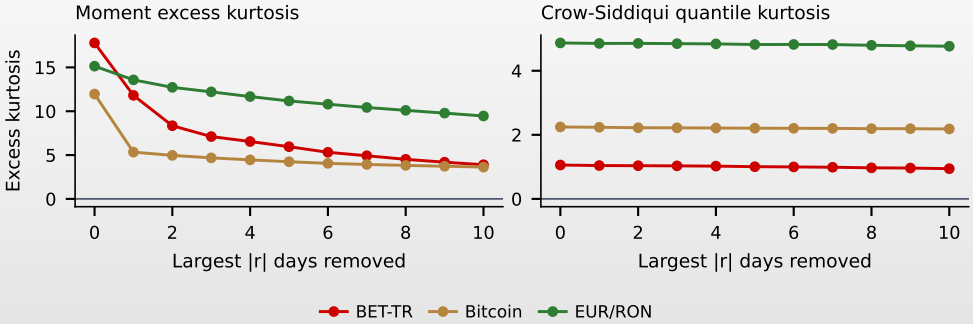
Back Exercise list

B1 [Solved]: BET-TR Step by Step

- Step 1: $\bar{r} = 0.077\%$, $\hat{\sigma} = 0.978\%$, volatility $\sqrt{252} \times 0.978\% = 15.5\%$
  - $\hat{S} = -1.40$, $\hat{K} - 3 = 17.79$
  - worst day -11.86% on 19 Dec 2018 (the bank-asset tax)
  - best day $+6.71\%$ on 24 Dec 2018
- Step 2: $JB = 40\,433$
  - skewness term $2\,992 \times 1.40^{\mathbf{2}}/6 = 977$
  - kurtosis term $2\,992 \times 17.79^{\mathbf{2}}/24 = 39\,455$, 98% of the statistic
- Step 3: $40\,433 > 5.99$, $p < 10^{-\mathbf{10}}$
- Step 4: excess kurtosis 17.8 with all days
  - 11.8 without 19 Dec 2018
  - 5.95 without the five largest days: 19 Dec 2018, 16 Mar 2020, 9 Mar 2020, 24 Dec 2018, 9 Mar 2022
- Step 5: Crow–Siddiqui 1.06 with all days, 1.01 without the five largest days
- One day carries a third of the moment kurtosis
  - the quantile measure does not move

	BET-TR	Bitcoin	EUR/RON
Mean (% a day)	0.077	0.118	0.007
Std. dev. (% a day)	0.978	3.485	0.279
Volatility (% a year)	15.5	66.6	4.4
Skewness	−1.40	−0.70	0.73
Excess kurtosis	17.8	12.0	15.1
Min (%)	−11.9	−46.5	−2.5
Max (%)	6.7	22.5	2.9
JB	40 433	26 482	51 575
Top 5 removed	5.95	4.24	11.17

B1 [Solved]: Sensitivity



- All JB statistics exceed the critical value 5.99 more than 4 000-fold
- Bitcoin without 12 Mar 2020 alone: kurtosis 12.0 → 5.3
- Interpretation: the tails are heavy by both measures
  - report the quantile excess kurtosis: its value does not hinge on a few days

MFM_ch1_seminar_tests

A3 [Solved]: Jarque–Bera and Ljung–Box, Step by Step — Task

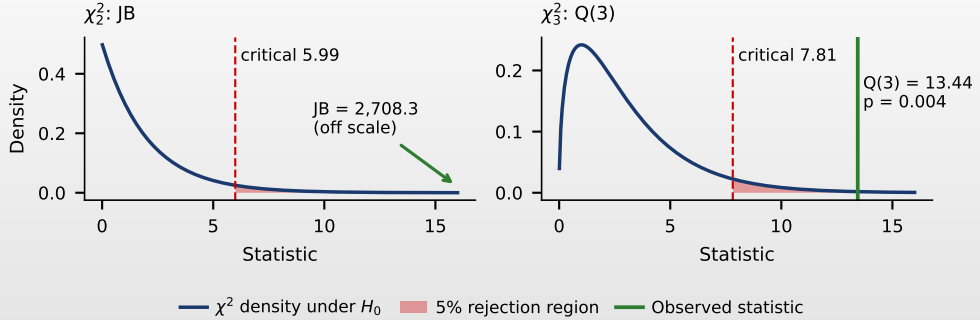
- **Questions**
  - how does a statistic become a decision?
  - which assumptions does the decision need?
- **Inputs**: two hypothetical samples of daily returns; tests at the 5% level
  - sample 1: $T = 2\,500$, skewness $\hat{S} = -0.5$, kurtosis $\hat{K} = 8$ (not excess)
  - sample 2: $T = 1\,000$, $\hat{\rho}_1 = 0.10$, $\hat{\rho}_2 = -0.05$, $\hat{\rho}_3 = 0.03$
- **Tasks**
  1. Compute the two terms of the JB statistic Jarque & Bera (1980) (definition), the statistic, and $p = P(\chi^2_2 \geq JB)$.
  2. Compute the three terms of the Ljung–Box $Q(3)$ Ljung & Box (1978) (definition, with $m = 3$), the statistic, and $p = P(\chi^2_3 \geq Q)$.
  3. State the H_0 of each test.
  4. Name the assumption of each χ^2 reference distribution that fails under volatility clustering.
- **Report**: the summands, the two statistics, the p-values and the decisions
- Notebook: section A3

Back Exercise list

A3 [Solved]: Step by Step

1. $\hat{S}^2/6 = 0.25/6 = 0.042$; $(\hat{K}-3)^2/24 = 25/24 = 1.042$
   - $JB = 2\,500 \times 1.084 = 2\,708.3$, $p < 10^{-10}$: reject
   - 96% of JB is the kurtosis term
2. $0.01/999 = 1.001 \times 10^{-5}$, $0.0025/998 = 0.251 \times 10^{-5}$, $0.0009/997 = 0.090 \times 10^{-5}$
   - $Q(3) = 1\,000 \times 1\,002 \times 1.342 \times 10^{-5} = 13.44$, $p = 0.004$: reject
3. H_0 of each test
   - JB: i.i.d. returns with the Normal distribution
   - Ljung–Box: no autocorrelation
4. The assumptions behind the reference distributions
   - Ljung–Box: χ^2_3 needs i.i.d. returns; under volatility clustering $\mathrm{Var}(\hat{\rho}_k) > 1/T$ and Q rejects too often (A6, B4)
   - JB: χ^2_2 also assumes independence; under volatility clustering a rejection concerns the joint "i.i.d. and Normal" null

A3 [Solved]: Reference Distributions



- The p-value is the area of the χ^2 density beyond the observed statistic
- Interpretation: for an i.i.d. sample both rejections are sound
  - for clustered returns only the JB rejection of "i.i.d. and Normal" is sound
  - the $Q(3)$ p-value is then too small

MFM_ch1_seminar_tests

B2 [Proposed]: Heavy Tails and VaR 1% on BET-TR — Task (1/2)

- **Model**: B1 [Solved]; notebook section B2
- **Question**: how large is the one-day VaR 1%, the loss threshold exceeded with probability 1%, under three distributions?
- **Data**: `r_bet`, $T = 2\,992$
  - models fitted and checked on the **same** sample: a descriptive comparison
  - not a backtest (a check of VaR on data after the estimation window)
- **Tasks**
  1. Standardise $z_t = (r_t - \bar{r})/\hat{\sigma}$ and count the days with $|z_t| > 3$.
  2. Compare their share with the share under the Normal distribution, $P(|Z| > 3) = 0.27\%$ for $Z \sim N(0,1)$.
  3. Fit a Student-t to z_t by maximum likelihood (`scipy.stats.t.fit`: degrees of freedom ν, location m, scale s).

Back Exercise list

B2 [Proposed]: Heavy Tails and VaR 1% on BET-TR — Task (2/2)

- **Tasks** (continued)
  4. Compute $\text{VaR}_{1\%} = -q_r$, with $q_r = \bar{r} + \hat{\sigma} q_z$, for three choices of the 1% quantile.
     - q_z: the 1% quantile of z_t; q_r: the same quantile in return units (undoing the standardisation)
     - Student-t: $q_z = m + s\, t_\nu^{-1}(0.01)$, with t_ν^{-1} the standard Student-t quantile
     - Normal distribution: $q_z = \Phi^{-1}(0.01) = -2.326$
     - empirical: $q_r =$ the 1% quantile of r_t
  5. Count the breaches (days with $r_t < q_r$, i.e. a loss larger than the VaR) and compare them with the expected count $0.01\, T = 29.9$.
- **Report**: a table method / VaR / breaches and a lower-tail plot with the three quantiles
- **Interpretation**: why does the empirical VaR have about 30 breaches by construction?

A4 [Solved]: Kurtosis of a Volatility Mixture — Task

- **Question**: can random volatility alone create heavy tails?
- **Assumptions**: $r_t = \sigma_t \varepsilon_t$, ε_t i.i.d. $N(0,1)$ and independent of σ_t
  - σ_t: the random volatility of day t; ε_t: the standardised shock
  - $\sigma_t = 1\%$ with probability 0.8 and 2% with probability 0.2
  - σ_t is drawn **independently every day**
- **Tasks**
  1. Using independence and $\mathbb{E}\varepsilon^4 = 3$, show that the population kurtosis (not the excess) is $K = \mathbb{E}r^4/(\mathbb{E}r^2)^2 = 3\,\mathbb{E}\sigma^4/(\mathbb{E}\sigma^2)^2$.
  2. Complete a two-state table with the probability, σ^2 (in $\%^2$) and σ^4 (in $\%^4$).
  3. Compute $\mathbb{E}\sigma^2$, $\mathbb{E}\sigma^4$ and K.
  4. Use Jensen's inequality, $\mathbb{E}[X^2] \geq (\mathbb{E}X)^2$ with $X = \sigma^2$, to show that $K \geq 3$.
  5. State whether this model produces volatility clustering.
- **Report**: the table, K and the excess kurtosis
- **Interpretation**: link the result to the mixture-of-distributions hypothesis of Clark (1973)
- Notebook: section A4

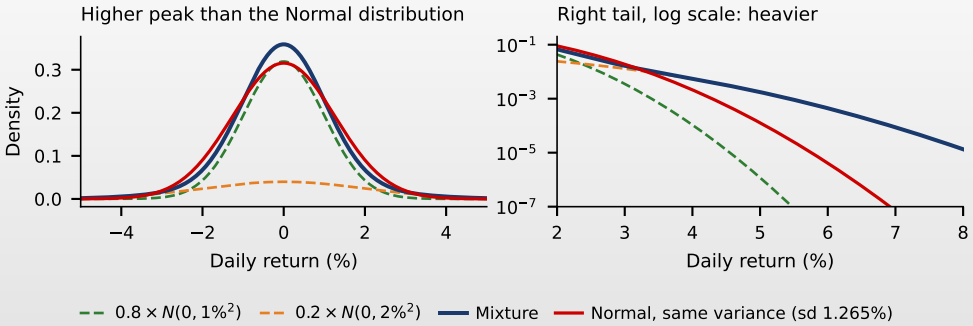
Back Exercise list

A4 [Solved]: Moments

State	Prob.	σ^2	σ^4
calm	0.8	1	1
turbulent	0.2	4	16
mean		1.6	4.0

1. $\mathbb{E}r^4 = \mathbb{E}\sigma^4\,\mathbb{E}\varepsilon^4 = 3\,\mathbb{E}\sigma^4$ and $\mathbb{E}r^2 = \mathbb{E}\sigma^2$
2. The table on the left
3. $\mathbb{E}\sigma^2 = 1.6$, $\mathbb{E}\sigma^4 = 4.0$
   - $K = 3 \times 4.0/1.6^2 = 12/2.56 = 4.69$; excess kurtosis 1.69
4. $\mathbb{E}[(\sigma^2)^2] \geq (\mathbb{E}\sigma^2)^2$, with equality only for a constant σ: $K \geq 3$
5. No clustering: the states are independent across days
   - so $|r_t|$ is uncorrelated
   - A9 makes the states persistent

A4 [Solved]: Densities



- $P(|r_t| > 4\%)$: 0.92% in the mixture, 0.16% for the Normal distribution with the same variance
- Interpretation: the random operational time of Clark (1973) (the pace of information arrival) is such a mixture
  - variation in volatility creates heavy tails

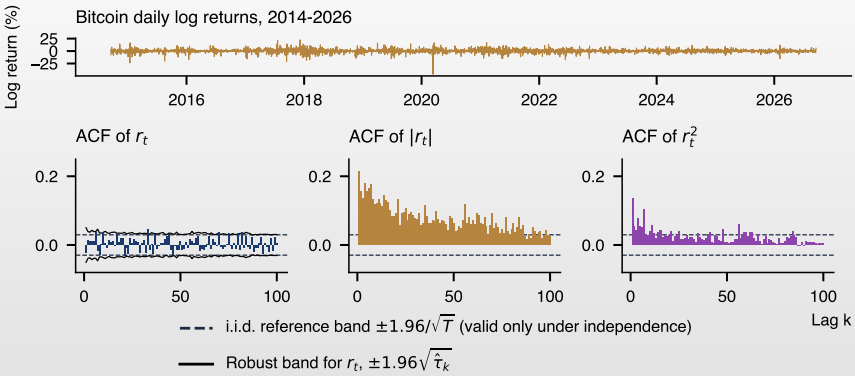
MFM_ch1_seminar_tests

B3 [Solved]: Uncorrelated but Not Independent (Bitcoin) — Task

- **Question**: do past Bitcoin returns predict the direction or only the size of today's return?
- **Data**: `r_btc`, 18 Sep 2014 – 18 Sep 2026, $T = 4\,384$
- **Tasks**
  1. Plot the returns.
  2. Plot the ACF of r_t, $|r_t|$ and r_t^2 for lags 1–100 with the i.i.d. reference band $\pm 1.96/\sqrt{T} = \pm 0.030$.
  3. Compute the Ljung–Box $Q(10)$ on r_t and on r_t^2 (`acorr_ljungbox`), with p-values from χ^2_{10}.
  4. Run the ARCH-LM test with $q = 5$ lags Engle (1982) (definition) and compare $(T-5)R^2$ with χ^2_5.
- **Report**: the chart, the lag-1 autocorrelations and the three statistics
  - label each with its series and its H_0
- **Interpretation**: the $Q(10)$ on r_t has $p \approx 3\%$
  - given step 4, is its i.i.d. reference distribution valid?
- Notebook: section B3

Back Exercise list

B3 [Solved]: Returns and ACF



- Direction: no structure in the ACF of r_t
- Size: $|r_t|$ and r_t^2 autocorrelated for dozens of lags

MFM_ch1_seminar_tests

B3 [Solved]: Tests

2. $\hat{\rho}_1$: -0.022 for r_t (inside the band), 0.213 for $|r_t|$ and 0.137 for r_t^2
3. $Q(10)$ results
   - on r_t: 20.2, $p = 0.028$
   - on r_t^2: 212.7, $p \approx 4 \times 10^{-40}$
4. ARCH-LM(5) $= 114.7$, $p \approx 4 \times 10^{-23}$
   - its χ^2 reference needs $\mathbb{E}r^4 < \infty$, so the p-value is indicative

⊡ The sizes of returns are dependent: returns are not i.i.d.
- the 2.8% p-value of $Q(10)$ on r_t rests on a false assumption

⊡ The i.i.d. band is a reference, not a valid test for $|r_t|$ and r_t^2
- A6 and B4 correct the inference for r_t

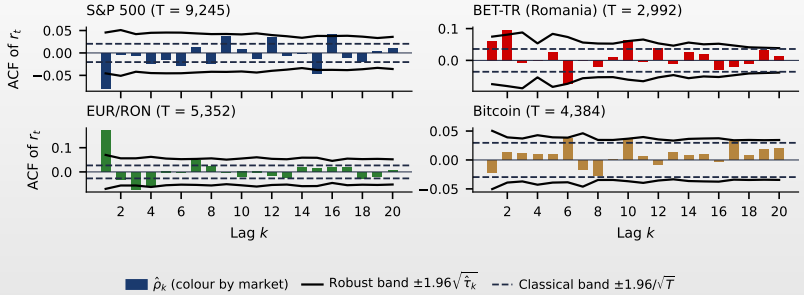
B4 [Solved]: Heteroskedasticity-Robust Ljung–Box — Task (1/2)

- **Question**: which autocorrelations of returns survive once the variance of $\hat{\rho}_k$ allows for clustering?
- **Data**: `r_sp`, `r_bet`, `r_eur`, `r_btc`; $x_t = r_t - \bar{r}$
- **Tasks**
  1. For $k = 1, \ldots, 20$ compute $\hat{\rho}_k$ and $\hat{\tau}_k$ Lobato, Nankervis & Savin (2001), the estimated variance of $\hat{\rho}_k$ (definition).
  2. Plot $\hat{\rho}_k$ with the classical band $\pm 1.96/\sqrt{T}$ and the robust band $\pm 1.96\sqrt{\hat{\tau}_k}$.
  3. Compute the classical $Q(10)$ and the robust $\tilde{Q}(10)$.
  4. Compare both with χ^2_{10}, whose 5% critical value is 18.3.
  5. Count the lags among 1–20 outside each band.
     - these are pointwise checks, distinct from the joint 10-lag test of steps 3–4

B4 [Solved]: Heteroskedasticity-Robust Ljung–Box — Task (2/2)

- **Assumptions of $\tilde{Q}$**
  - a finite fourth moment, $\mathbb{E}r^4 < \infty$, needed for $\hat{\tau}_k$ to estimate a finite variance
  - zero cross terms, $\mathbb{E}[x_t^2 x_{t-k} x_{t-l}] = 0$ for $k \neq l$
  - i.e. x_t^2 is uncorrelated with the products $x_{t-k} x_{t-l}$ of past returns at different lags
- **Interpretation**
  - for which series does autocorrelation survive?
  - is the fourth-moment assumption credible?
- Notebook: section B4

B4 [Solved]: ACF with Both Bands



- S&P 500, lag 1: $\hat{\rho}_{\mathbf{1}} = -0.079$, $T\hat{\tau}_{\mathbf{1}} = 5.00$
  - the variance of $\hat{\rho}_{\mathbf{1}}$ is 5 times the i.i.d. value $1/T$
  - the robust band is $1.96\sqrt{5.00/9\,245} = 0.046$ instead of 0.020
- At lag 1 the robust band is $\sqrt{T\hat{\tau}_{\mathbf{1}}} = 1.7$ (Bitcoin) to 2.6 (EUR/RON) times wider than the classical one

MFM_ch1_seminar_tests

B4 [Solved]: Joint Tests and Interpretation

	S&P 500	BET-TR	EUR/RON	Bitcoin
$Q(10)$ (p)	91.6 ($< 10^{-10}$)	70.0 ($< 10^{-10}$)	221.7 ($< 10^{-10}$)	20.2 (0.028)
$\tilde{Q}(10)$ (p)	19.0 (0.041)	16.9 (0.076)	37.4 (< 0.001)	11.9 (0.29)
Lags outside, classical / robust	8 / 3	5 / 2	6 / 2	3 / 0
Lags outside the robust band	1, 15, 16	2, 10	1, 3	none

- At 5% the classical test rejects for all four series
  - the robust test rejects for the S&P 500 (barely) and EUR/RON only
- Bitcoin (B3): the 2.8% p-value becomes 29%
  - no evidence of linear autocorrelation at lags 1–10
- EUR/RON keeps $\hat{\rho}_1 = 0.17$
  - under the managed float, the BNR reference rate adjusts in small steps over several days

MFM_ch1_seminar_tests

B4 [Solved]: Remaining Assumptions

- Zero cross terms
  - the test of Dalla, Giraitis & Phillips (2022) keeps only the significant cross terms
  - it finds none for the S&P 500, BET-TR and Bitcoin
- Finite fourth moment
  - if $P(|r_t| > x)$ decays like $x^{-\alpha}$ with a tail index $\alpha \leq 4$, $\mathbb{E}r^4$ is infinite
  - the robust p-values are then not valid
  - A5 and B6 estimate this tail index after the lecture

B5 [Proposed]: Aggregational Gaussianity (Bitcoin) — Task

- **Model**: B1 and B3 [Solved]; notebook section B5
- **Question**: do h-day Bitcoin returns approach the Normal distribution as h grows?
- **Data**: `r_btc`, $T = 4\,384$, in non-overlapping blocks from the first day
  - the first 5-day block is 18–22 Sep 2014, the second 23–27 Sep 2014
  - the incomplete last block (4 days for $h = 5$ and $h = 20$) is dropped
- **Tasks**
  1. Build the h-day log returns (sums of h consecutive daily returns) for $h = 1, 5, 20$.
  2. Draw a QQ plot of each standardised series against the Normal distribution, with N, the number of blocks, in the title.
  3. Compute the excess kurtosis and the i.i.d.-Gaussian reference SE $\sqrt{24/N}$.
     - $\sqrt{24/N}$: the standard error of the excess kurtosis estimated from N i.i.d. Normal observations (24, as in JB)
- **Report**: the three QQ plots and a table h, N, excess kurtosis, $\sqrt{24/N}$
- **Interpretation**
  - why is $\sqrt{24/N}$ too optimistic here?
  - why do non-overlapping blocks remove shared days but not volatility dependence?

Back Exercise list

C1 [Proposed]: Critique an AI Answer — the Answer

- **Prompt** sent to an AI assistant: "Are the daily BET-TR log returns Normally distributed? Are they autocorrelated? Give Python code and a conclusion"
- **Answer**: assume NumPy is imported as `np` and `bet` has a `close` column; assess changes to numbers or conclusions

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

- **Conclusion of the AI**
  - "Excess kurtosis 14.8: heavy tails"
  - "JB p-value 0.000: the probability that BET-TR returns follow the Normal distribution is 0%"
  - "Ljung–Box $Q(10) = 713.5$, $p < 0.001$: returns are strongly autocorrelated, so BET-TR is predictable and not a random walk"

C1 [Proposed]: Critique an AI Answer — Task (1/2)

- **Model**: A3, B3 and B4 [Solved]; notebook section C1
- **Data**: `r_bet`, the 2 992 returns of the setup
- **Tasks**
  1. Find every substantive error in the code and in the three conclusions.
  2. For each error, say how you would detect it.
     - a test on simulated data or a number from the primer slides
     - a solved exercise or the library documentation
  3. Write corrected, executable code.

C1 [Proposed]: Critique an AI Answer — Task (2/2)

- **Tasks** (continued)
  4. Report the excess kurtosis, JB with its p-value, $Q(10)$ on r_t and on r_t^2, and the robust $\tilde{Q}(10)$ of B4.
  5. State the assumptions behind the robust p-value.
  6. Write the corrected conclusion in three sentences.
- **Report**
  - the corrected code
  - a table AI value / correct value
  - a chart of the ACF of r_t and of r_t^2
- AI code usually runs without an error message
  - the mistakes are in the statistics, not in the syntax
  - log what you find in `AI_ERRORS.md`

A5 [Solved]: The Hill Estimator, Step by Step — Task (1/2)

- **Questions**
  - how is a tail index estimated from the largest observations?
  - how precise is it?
- **Inputs**: the 11 largest absolute daily returns (%) of a hypothetical sample: 12.0; 9.5; 8.2; 7.4; 6.9; 6.1; 5.8; 5.5; 5.2; 5.0; 4.8
- **Notation**
  - $X_{(1)} \geq X_{(2)} \geq \ldots$: the absolute values sorted in decreasing order (the order statistics)
  - k: the number of tail observations; $X_{(k+1)}$: the threshold where the tail starts
- **Assumption**: above the threshold $X_{(k+1)}$ the tail is Pareto, $P(X > x) \propto x^{-\alpha}$
  - α: the tail index; a small α means a heavy tail; $\mathbb{E}|X|^p$ is finite only for $p < \alpha$
  - the observations are independent

Back Exercise list

A5 [Solved]: The Hill Estimator, Step by Step — Task (2/2)

- **Hill estimator** Hill (1975): the inverse of the mean log-ratio between the k largest values and the threshold

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

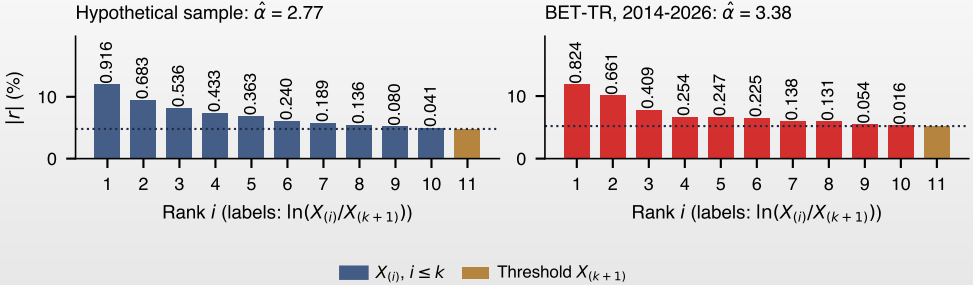
  - a heavy tail has large log-ratios, hence a small $\hat{\alpha}$

- **Tasks**
  1. For $k = 10$ compute the log-ratios $\ln(X_{(i)}/X_{(11)})$, their mean and $\hat{\alpha} = 1/\text{mean}$.
  2. Compute the standard error $\text{SE} = \hat{\alpha}/\sqrt{k}$ (valid for i.i.d. observations) and the 95% interval $\hat{\alpha} \pm 1.96\,\text{SE}$.
  3. State which moments $\mathbb{E}|X|^p$ ($p < \alpha$) exist at the point estimate.
  4. State which of these claims hold at 95%.
  5. Repeat steps 1–2 on the 11 largest $|r_t|$ of BET-TR, using the worksheet printed by the notebook.

- **Report**: $\hat{\alpha}$, SE and interval for both samples

- Notebook: section A5

A5 [Solved]: Hypothetical Sample



1. The labels are the log-ratios to 4.8: sum 3.616, mean 0.3616, $\hat{\alpha} = 1/0.3616 = 2.77$
2. SE $= 2.77/\sqrt{10} = 0.87$; interval [1.05; 4.48]
   - a rough approximation by the Normal distribution, valid for i.i.d. Pareto tails
3. At 2.77 the variance exists and the fourth moment does not
4. The interval contains values below 2 and above 4: neither claim holds at 95%

MFM_ch1_seminar_tests

A5 [Solved]: BET-TR Worksheet

i	Date	r_t (%)	$X_{(i)}$ (%)	$\ln(X_{(i)}/X_{(\mathbf{11})})$
1	19 Dec 2018	-11.857	11.857	0.824
2	16 Mar 2020	-10.075	10.075	0.661
3	9 Mar 2020	-7.827	7.827	0.409
4	24 Dec 2018	$+6.707$	6.707	0.254
5	9 Mar 2022	$+6.662$	6.662	0.247
6	24 Aug 2015	-6.512	6.512	0.225
7	24 Mar 2020	$+5.972$	5.972	0.138
8	17 Mar 2020	$+5.928$	5.928	0.131
9	4 Mar 2022	-5.490	5.490	0.054
10	12 Mar 2020	-5.285	5.285	0.016
11	18 Jan 2016	-5.202	5.202	threshold

- Row 1 worked: $\ln(11.857/5.202) = \ln 2.279 = 0.824$
- Sum of the ten log-ratios 2.958; $\hat{\alpha}_{\mathbf{10}} = 10/2.958 = 3.38$
- SE $= 3.38/\sqrt{10} = 1.07$; interval [1.28; 5.48]
- Seven of the ten days fall in December 2018 or March 2020
  - the extremes cluster, so the i.i.d. SE understates the uncertainty
- Interpretation: ten observations cannot decide whether the fourth moment exists
  - B6 uses $k = 2\%$ of n

B6 [Solved]: Hill on Both Tails, BET-TR and Bitcoin — Task (1/2)

- **Question**: are large losses more frequent than large gains of the same size?
- **Data**: `r_bet` ($n = 2\,992$), `r_btc` ($n = 4\,384$)
  - losses `loss = -r[r < 0]`, gains `gain = r[r > 0]`
  - each sorted in decreasing order
- **Notation**
  - $\hat{\alpha}(k)$: the Hill estimator of A5, computed from the k largest values of one tail
  - $\hat{\alpha}_-$: the tail index of losses; $\hat{\alpha}_+$: that of gains; $\hat{\alpha}_- < \hat{\alpha}_+$ means losses have the heavier tail
  - $\lfloor \cdot \rfloor$: the integer part
- **Tasks**
  1. Plot the Hill estimate $\hat{\alpha}(k)$ of each tail for k from 1% to 10% of the full-sample n, with the band $\hat{\alpha}(1 \pm 1.96/\sqrt{k})$.
  2. Compute $\hat{\alpha}_-$ and $\hat{\alpha}_+$ at $k = \lfloor 0.02\, n \rfloor$ (BET-TR 59, Bitcoin 87), with 95% intervals.
     - the lecture uses $k = 2.5\%$ of n for $|r_t|$, which pools both tails

Back Exercise list

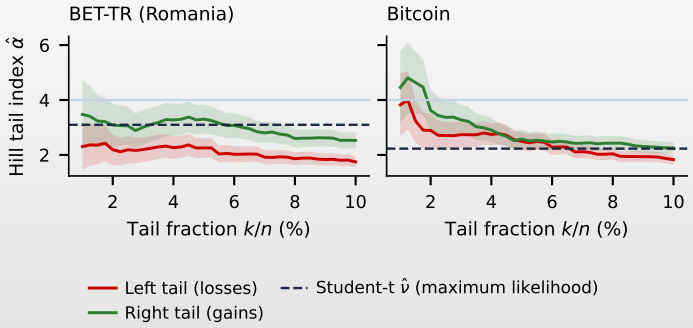
B6 [Solved]: Hill on Both Tails, BET-TR and Bitcoin — Task (2/2)

- **Tasks** (continued)
  3. Compute the approximate independent-tail statistic and its two-sided p-value.

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

     - denominator: the standard error of the difference, from $\mathrm{SE} = \hat{\alpha}/\sqrt{k}$ for each tail
     - under $H_0 : \alpha_- = \alpha_+$, $z \approx N(0,1)$; $|z| > 1.96$: reject H_0 at 5%
  4. Compare with the Student-t $\hat{\nu}$ fitted to the standardised returns: 3.10 for BET-TR (B2), 2.23 for Bitcoin.
- **Interpretation**
  - is the gain/loss asymmetry statistically established?
  - how does the dependence between the two tail estimators change the SE of the difference?
- Notebook: section B6

B6 [Solved]: Hill Plots and Estimates



- BET-TR: $\hat{\alpha}_- = 2.19$ [1.63; 2.75], $\hat{\alpha}_+ = 3.08$ [2.30; 3.87]
- Bitcoin: $\hat{\alpha}_- = 2.89$ [2.28; 3.50], $\hat{\alpha}_+ = 3.62$ [2.86; 4.38]
- E.g. BET-TR losses: $2.19 \times (1 \pm 1.96/\sqrt{59}) = 2.19 \times (1 \pm 0.255)$

MFM_ch1_seminar_tests

B6 [Solved]: Test and Interpretation

- BET-TR: $\sqrt{2.19^2/59 + 3.08^2/59} = \sqrt{0.0812 + 0.1612} = 0.492$; $z = (2.19 - 3.08)/0.492 = -1.82$, $p = 0.069$
- Bitcoin: $\sqrt{2.89^2/87 + 3.62^2/87} = 0.497$; $z = (2.89 - 3.62)/0.497 = -1.47$, $p = 0.14$
- Decision: losses are heavier in both markets, but $H_0 : \alpha_- = \alpha_+$ is not rejected at 5%
  - the statistic treats the two tails as independent samples
  - dependence changes the uncertainty: $\mathrm{Var}(\hat{\alpha}_- - \hat{\alpha}_+) = \mathrm{Var}(\hat{\alpha}_-) + \mathrm{Var}(\hat{\alpha}_+) - 2\,\mathrm{Cov}(\hat{\alpha}_-, \hat{\alpha}_+)$
  - clustering of extremes can raise each variance, and the covariance term can offset it: the direction is not known in advance
  - a joint block bootstrap of the original returns estimates its direction and size (extension)
- $\hat{\nu}$ is fitted to the whole distribution, while $\hat{\alpha}$ uses only the tail
  - Bitcoin's 2.23 lies below both tail indices
  - the Student-t does not fit centre and tails together

MFM_ch1_seminar_tests

A6 [Proposed]: Widening of the Autocorrelation Band under Clustering — Task (1/2)

- **Model**: A4 [Solved]; notebook section A6
- **Setting**: $x_t = \sigma_t \varepsilon_t$ with ε_t i.i.d. $N(0,1)$, independent of the stationary volatility process
  - then x_t is uncorrelated
  - $\mathrm{Var}(\hat{\rho}_1) \approx v_1/T$, with $v_1 = \mathbb{E}[x_t^2 x_{t-1}^2]/(\mathbb{E}x_t^2)^2$ Lobato, Nankervis & Savin (2001)
  - v_1: the factor by which volatility clustering inflates the variance of $\hat{\rho}_1$ relative to the i.i.d. value $1/T$
- **Inputs**
  - A4: $\mathbb{E}\sigma^2 = 1.6$, $\mathrm{Var}(\sigma^2) = 4.0 - 1.6^2 = 1.44$ (in $\%^4$)
  - persistent regimes (A9): $\mathrm{Cov}(\sigma_t^2, \sigma_{t-1}^2) = \lambda \mathrm{Var}(\sigma^2)$ with $\lambda = 0.875$
  - $\lambda \in [0,1)$: the persistence of the regimes; $\lambda = 0$: states independent across days
  - S&P 500: $T = 9\,245$, $\hat{\rho}_1 = -0.079$, $T\hat{\tau}_1 = 5.00$

A6 [Proposed]: Widening of the Autocorrelation Band under Clustering — Task (2/2)

- **Tasks**
  1. Show that $v_1 = 1$ for the A4 mixture, whose states are independent across days.
  2. Show that $v_1 = 1 + \lambda\,\mathrm{Var}(\sigma^2)/(\mathbb{E}\sigma^2)^2$ with persistent regimes.
  3. Compute v_1 and the band-widening factor $\sqrt{v_1}$.
  4. Compare with B4: compute the S&P 500 lag-1 statistic $t = \hat{\rho}_1/\mathrm{SE}$ with the i.i.d. SE $1/\sqrt{T}$ and with the robust SE $\sqrt{\hat{\tau}_1}$ ($|t| > 1.96$: significant at 5%).
  5. Compute the two 95% band half-widths.
- **Interpretation**: which feature widens the band: heavy tails or persistent volatility?
- **Report**: v_1 and $\sqrt{v_1}$; a two-row i.i.d./robust table of t statistics and band half-widths

After the Lecture: Remaining Tasks

- These three tasks use tools that Chapter 1 introduces
  - A5 [Solved] and B6 [Solved]: the Hill estimator of the tail index, step by step and on both tails
  - A6 [Proposed]: which mechanism changes the autocorrelation band in B4
- Work through A5 and B6 after the lecture
- Solve A6 before the next seminar; its solution is discussed then

Extension Packs: Content and Prerequisites

Pack	Exercises	Prerequisites
1. Return arithmetic and drawdowns	A1(c) [Solved], A2(b–f), A7	A1, A2(a)
2. Mixtures, regimes and tails	A8 [Solved], A9, A10, A11	A4, A5, A6
3. Robust moments and model checks	B1(d) [Solved], B7, B8	B1, B4, B6
4. Dependence and asymmetry	B9, B10, B11, B12	B3, B4
5. Performance and range estimators	B13, B14, B15	A1, A8, B1
6. Projects	C2, C3	B4, the lecture case study

- Notebook sections carry the same labels as the exercises
- Solutions of [Proposed] exercises are discussed at the next seminar

The Random Walk and Brownian Motion

- **Random walk**: the log price, $X_t = \ln P_t$, adds up independent shocks

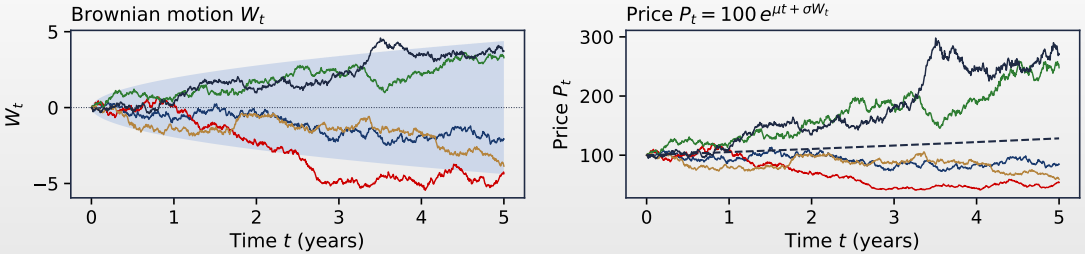
$$X_t = X_{t-1} + r_t, \qquad r_t \text{ i.i.d. } N(\mu\Delta,\ \sigma^2\Delta)$$

  - Δ: the length of one step, in years (one trading day: $\Delta = 1/252$)
  - μ: the annual drift of X_t; σ: the annual volatility
  - the benchmark against which the stylised facts are defined: i.i.d. returns with the Normal distribution
- **Standard Brownian motion** W_t, $t \geq 0$: the limit of the random walk as $\Delta \to 0$
  - $W_0 = 0$; the increment $W_t - W_s$, $s < t$, is independent of the past and has the distribution $N(0,\ t-s)$
  - hence $W_t \sim N(0, t)$: the standard deviation grows like $\sqrt{t}$, and $|W_t| \leq 1.96\sqrt{t}$ with probability 95%
- **The price** as geometric Brownian motion (A2, A8)

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

  - $P_t > 0$; the log return over an interval of length Δ has the distribution $N(\mu\Delta,\ \sigma^2\Delta)$

Simulated Brownian Paths



- Left: five simulated paths of W_t over 5 years and the 95% band $\pm 1.96\sqrt{t}$, which widens like $\sqrt{t}$
- Right: the same paths turned into prices, $P_t = 100\,e^{\mu t + \sigma W_t}$, with $\mu = 0.05$ and $\sigma = 0.20$
  - after 5 years the simulated prices lie between 54 and 270: uncertainty grows with the horizon
  - the dashed line $100\,e^{\mu t}$ is the median price, because the median of W_t is 0

A1(c) [Solved]: Portfolio Returns — Task and Solution

- **Question**: which return adds up across assets?
- **Inputs**: 60% in asset A and 40% in asset B for one day; $R_A = +5\%$, $R_B = -2\%$
- **Tasks**
  1. Compute the portfolio simple return $R_p = 0.6R_A + 0.4R_B$.
  2. Compute the weighted log return $0.6\, r_A + 0.4\, r_B$, with $r_A = \ln(1 + R_A)$, $r_B = \ln(1 + R_B)$.
  3. Compute the true portfolio log return $\ln(1 + R_p)$.
- **Solution**
  1. $R_p = 0.6 \times 5\% + 0.4 \times (-2\%) = 2.20\%$
  2. $0.6 \times \ln 1.05 + 0.4 \times \ln 0.98 = 0.6 \times 4.879\% - 0.4 \times 2.020\% = 2.119\%$
  3. $\ln 1.022 = 2.176\%$
- Interpretation: simple returns add across assets, log returns add over time (A1)
  - for daily moves the gap between the two rules is small
  - for large moves it is not

Back Exercise list

A2(b–f) [Proposed]: A Benchmark for Drawdowns — Task (1/2)

- **Model**: A2(a); notebook section A2
- **Assumption**: the log price is $X_t = \mu t + \sigma W_t$
  - W_t a standard Brownian motion: independent increments, $W_t \sim N(0, t)$
  - μ: the annual drift of the log price; σ: the annual volatility; T: the horizon, in years (here not the number of returns)
  - maximum log drawdown $D = \max_{s \le t \le T}(X_s - X_t) = -\ln(1 + MDD)$: the largest fall of X from a peak at s to a later time t
  - for $\mu = 0$: $\mathbb{E}[D] = \sqrt{\pi/2}\,\sigma\sqrt{T}$ Magdon-Ismail et al. (2004), with $\sqrt{\pi/2} = 1.2533$
- **Inputs**: gold spot, 1990–2026
  - $\sigma = 16.3\%$, $T = 36.7$ years, mean log return 6.5% a year
  - observed $MDD = -44.7\%$, so $D = 0.592$

Back Exercise list

A2(b–f) [Proposed]: A Benchmark for Drawdowns — Task (2/2)

- **Tasks**
  1. Use $W(cT) \overset{d}{=} \sqrt{c}\, W(T)$ to explain why $\mathbb{E}[D]$ is proportional to $\sqrt{T}$ when $\mu = 0$ ($c > 0$: the time-scaling factor; $\overset{d}{=}$: same distribution).
  2. Compute the ratio of $\mathbb{E}[D]$ for 40 and 10 years of the same process.
  3. Compute the driftless benchmark for gold and compare it with the observed D.
  4. Simulate 2 000 daily paths (`numpy.random.default_rng(42)`), starting at log price zero, with 252 steps a year and the gold volatility above with $\mu = 0$ and with $\mu = 6.5\%$.
  5. Simulate the 40-year paths once; evaluate their prefixes of `round(252*T)` steps (36.7 years: 9 248 steps). Plot and tabulate mean D against T for both drifts at each year 1–40 and 36.7 years.
- **Report**: the ratio from step 2, the gold benchmark $\mathbb{E}[D]$, and the table and chart of mean D against T
- **Interpretation**: what do steps 1–5 imply for Calmar ratios computed on samples of different length?

A7 [Proposed]: The $\sqrt{h}$ Rule under Autocorrelation — Task (1/2)

- **Model**: A1 [Solved]; notebook section A7
- **Assumption**: returns follow a stationary first-order autoregressive process, AR(1)

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

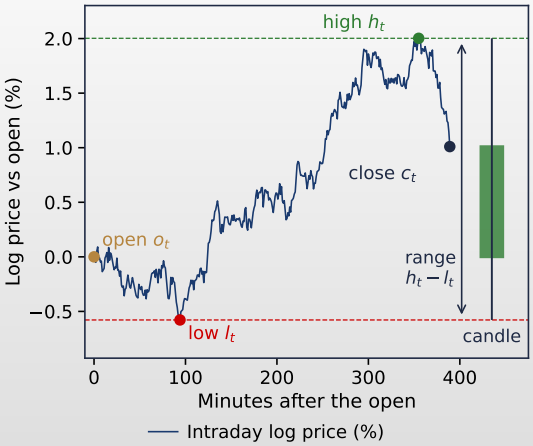
  - e_t: a shock uncorrelated over time, with mean 0
  - ρ: the autoregressive coefficient, equal to the lag-1 autocorrelation; $\rho > 0$: returns of the same sign tend to follow each other
  - $\mathrm{Var}(r_t) = \sigma^2$; the autocovariance at lag j is $\gamma_j = \mathrm{Cov}(r_t, r_{t-j}) = \sigma^2 \rho^j$, and $\gamma_0 = \sigma^2$
- **The $\sqrt{h}$ rule**: the variance of an h-day return is $h\sigma^2$, true only for uncorrelated returns
- **Variance ratio** (VR): $VR(h) = \mathrm{Var}\left(\sum_{j=1}^{h} r_{t+j}\right)/(h\sigma^2)$
  - $VR(h) = 1$ for uncorrelated returns; $VR(h) > 1$ for $\rho > 0$, so the $\sqrt{h}$ rule understates risk
- **Example**, $h = 3$: $\mathrm{Var}(r_1 + r_2 + r_3) = 3\gamma_0 + 2(2\gamma_1 + \gamma_2)$: two pairs at lag 1, one pair at lag 2

Back Exercise list

A7 [Proposed]: The $\sqrt{h}$ Rule under Autocorrelation — Task (2/2)

- **Tasks**
  1. Generalise the pair count: show that $\mathrm{Var}\left(\sum_{j=1}^{h} r_{t+j}\right) = h\sigma^2\left[1 + 2\sum_{j=1}^{h-1}(1 - j/h)\rho^j\right]$.
  2. Using $\sum_{j=1}^{h-1}\rho^j = \rho(1-\rho^{h-1})/(1-\rho)$ and $\sum_{j\geq 1} j|\rho|^j = |\rho|/(1-|\rho|)^2$, show that $VR(h) \to (1+\rho)/(1-\rho)$ as $h \to \infty$.
  3. Illustrative calculation: compute $\sqrt{VR(252)}$ and the corrected annual volatilities, treating the sample lag-1 autocorrelations as AR(1) parameters.
     - EUR/RON: $\rho = 0.170$, volatility 4.44%
     - S&P 500: $\rho = -0.079$, volatility 17.99%
  4. Explain why volatility clustering alone does not change step 1.
- **Interpretation**: which ACF check must precede step 3?
- **Report**: the variance derivation, two corrected annual volatilities and the ACF check.

The Daily Range and Parkinson Volatility



- For the extension packs (A8, B15)
- O_t, H_t, L_t, C_t: the open, high, low and close of day t
  - in lower case, their logs: $o_t = \ln O_t$, $h_t = \ln H_t$, $l_t = \ln L_t$, $c_t = \ln C_t$
- **Log range** $h_t - l_t = \ln(H_t/L_t)$: the amplitude of the day
  - it uses the high and the low of the path, not only the closing price
  - in the chart: the range is 2.58%, while $c_t - o_t$ is only 1.01%
- **Parkinson estimator** Parkinson (1980)

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

  - n: the number of days; $\hat{\sigma}_P^2$: the estimated daily variance
  - $4 \ln 2 = 2.77$: for a driftless Brownian motion, $\mathbb{E}[(h_t - l_t)^2] = 4 \ln 2 \, \sigma^2$

A8 [Solved]: Parkinson Volatility and Its Efficiency — Task (1/2)

- **Question**: how much information does the daily range carry about volatility?
- **Notation**
  - O_t, H_t, L_t, C_t: the open, high, low and close of day t
  - their logs: $o_t = \ln O_t,\ h_t = \ln H_t,\ l_t = \ln L_t,\ c_t = \ln C_t$
  - the log range: $h_t - l_t = \ln(H_t/L_t)$
- **Inputs**: H/L of five days: 1.020, 1.015, 1.030, 1.010, 1.025; $A = 252$
- **Assumption**: within the day the log price is a driftless Brownian motion with volatility σ
  - no overnight jumps
- **Parkinson estimator** Parkinson (1980): the daily variance estimated from the range

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

  - $\overline{(h-l)^2}$: the average over days of the squared range
  - $4\ln 2 = 2.7726$: under the assumption $\mathbb{E}[(h-l)^2] = 4\ln 2\,\sigma^2$, so the division makes the estimator unbiased

Back Exercise list

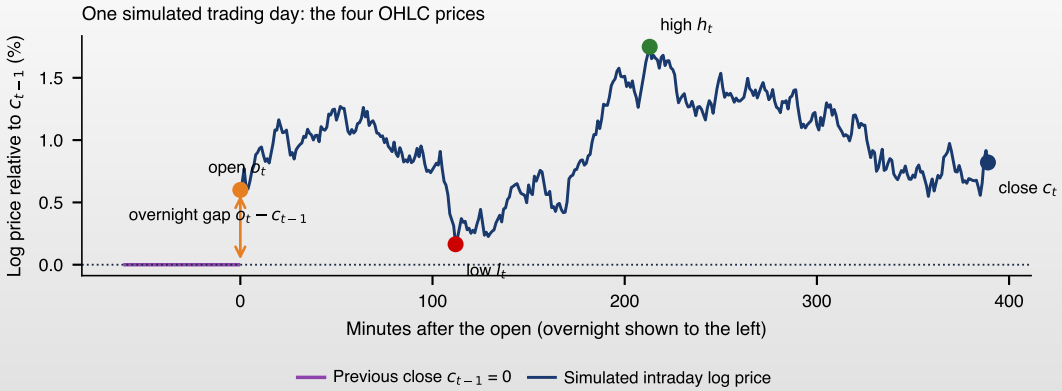
A8 [Solved]: Parkinson Volatility and Its Efficiency — Task (2/2)

Tasks

1. Complete a five-row table: $\ln(H/L)$, its square, and the mean of the squares.
2. Compute $\hat{\sigma}_P^2$, the daily and the annual volatility.
3. Name the information the estimator ignores.
4. Explain the effect of a strong drift on the estimate.
5. Using $\mathbb{E}[(h-l)^2] = 4\ln 2\,\sigma^2$ and $\mathbb{E}[(h-l)^4] = 9\zeta(3)\,\sigma^4$, compare the per-day variance of $(h-l)^2/(4\ln 2)$ with $\mathrm{Var}(r_t^2) = 2\sigma^4$.
   - $\zeta(3) = \sum_{n\geq 1} n^{-3} = 1.202$: the Riemann zeta function at 3
   - r_t^2: the close-to-close estimator of the daily variance, from a return with distribution $N(0, \sigma^2)$
   - the estimator with the smaller variance is more efficient: it needs fewer days for the same precision

Notebook: section A8

A8 [Solved]: One Simulated Trading Day



A8 [Solved]: Parkinson Estimate from Five Ranges

H/L	$\ln(H/L)$	square ($\times 10^{-4}$)
1.020	0.0198	3.92
1.015	0.0149	2.22
1.030	0.0296	8.74
1.010	0.0100	0.99
1.025	0.0247	6.10
mean		4.39

2. $\hat{\sigma}_P^2 = 0.000439/2.7726 = 0.000158$; $\hat{\sigma}_P = 1.26\%$ a day, $1.26\% \times 15.87 = 20.0\%$ a year
3. Ignored: the open and the overnight jump Yang & Zhang (2000)
4. A drift widens the range and biases $\hat{\sigma}_P$ up Rogers & Satchell (1991)

- A simulated day: the range $h_t - l_t$ uses the whole path, the close-to-close return only its end points

MFM_ch1_seminar_tests

A8 [Solved]: Efficiency

5. $\mathrm{Var}\big((h-l)^2\big) = \big(9\zeta(3) - 16(\ln 2)^2\big)\sigma^4 = (10.819 - 7.687)\,\sigma^4 = 3.131\,\sigma^4$
   - divided by $(4\ln 2)^2 = 7.687$: $0.407\,\sigma^4$ per day for the Parkinson term
   - close-to-close: $\mathrm{Var}(r_t^2) = 3\sigma^4 - \sigma^4 = 2\sigma^4$ per day

- For an n-day average of independent days, both variances are divided by n
  - the ratio stays $2/0.407 = 4.9$
- One day of ranges carries about the information of five days of close-to-close returns
- Interpretation: the gain assumes driftless continuous trading without overnight jumps
  - a stale open (an open computed from prices of components that have not yet traded) affects the estimators that use the open
  - namely Garman–Klass, Rogers–Satchell and Yang–Zhang, not Parkinson (B15)

A9 [Proposed]: The Origin of the Autocorrelation of r_t^2 — Task (1/2)

- **Model**: A4 [Solved] and A6; notebook section A9
- **Assumption**: as in A4, $\sigma_t^2 \in \{1, 4\}$ (in $\%^2$) with stationary probabilities (0.8; 0.2)
  - but the state follows a two-state Markov chain Hamilton (1989): tomorrow's state depends only on today's
  - transition matrix $\begin{pmatrix} 0.975 & 0.025 \\ 0.1 & 0.9 \end{pmatrix}$ (rows: today calm L, turbulent H)
  - $p_{LL} = 0.975$ and $p_{HH} = 0.9$ are the probabilities of staying in the same state
  - $\pi_L = 0.8$, $\pi_H = 0.2$: the stationary probabilities, i.e. the long-run shares of the two states
- **Tasks**
  1. Check the balance condition $\pi_L(1 - p_{LL}) = \pi_H(1 - p_{HH})$, i.e. $0.8 \times 0.025 = 0.2 \times 0.1$: the flow from L to H equals the flow from H to L.
  2. Compute $\lambda = p_{LL} + p_{HH} - 1$, the persistence of the chain: $\lambda = 0$ for independent states, $\lambda \to 1$ for very persistent regimes.

Back Exercise list

A9 [Proposed]: The Origin of the Autocorrelation of r_t^2 — Task (2/2)

- **Tasks** (continued)
  3. With the indicator $I_t = 1$ in the turbulent state, $\sigma_t^2 = 1 + 3I_t$; show that $\mathrm{Cov}(I_t, I_{t-k}) = \pi_H(1 - \pi_H)\lambda^k$.
  4. Deduce that $\mathrm{Cov}(r_t^2, r_{t-k}^2) = \mathrm{Var}(\sigma^2)\lambda^k$ for $k \geq 1$.
  5. Using A4, compute $\mathrm{Var}(\sigma^2)$, $\mathrm{Var}(r^2) = 3\mathbb{E}\sigma^4 - (\mathbb{E}\sigma^2)^2$ and $\mathrm{corr}(r_t^2, r_{t-k}^2)$ at lags 1 and 10.
  6. Plot the implied ACF against the S&P 500 ACF of r_t^2 (0.31 at lag 1, 0.04 at lag 100).
- **Interpretation**: which two features of the data can this model not reproduce?

A10 [Proposed]: The S&P 500 Tails in a GARCH with Normal Innovations — Task (1/2)

- **Model**: A4 [Solved]; notebook section A10
  - an advanced preview of Chapter 5
- **Model equations** GARCH(1,1) Bollerslev (1986): today's variance depends on yesterday's shock and yesterday's variance

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

  - μ: the mean return; ε_t: the shock of day t; z_t: the standardised innovation, i.i.d. $N(0,1)$
  - σ_t^2: the conditional variance, known at the end of day $t-1$
  - $\omega > 0$: the base level of the variance; $\alpha \geq 0$: the reaction to yesterday's shock; $\beta \geq 0$: the persistence of yesterday's variance
  - $\alpha + \beta < 1$: the variance reverts to its long-run mean $\omega/(1-\alpha-\beta)$
  - here α is a GARCH coefficient, not a tail index

Back Exercise list

A10 [Proposed]: The S&P 500 Tails in a GARCH with Normal Innovations — Task (2/2)

- **Result**: if $1-(\alpha+\beta)^2-2\alpha^2>0$, the unconditional kurtosis of r_t is

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

- **Tasks**
  1. Compute K for $(\alpha,\beta)=(0.10;\ 0.85)$ and $(0.10;\ 0.88)$.
  2. State the condition for a finite fourth moment.
  3. Compare it with $\alpha+\beta<1$.
  4. With $\alpha+\beta=0.99$, solve for the α that gives the S&P 500 kurtosis $K=13.9$.
  5. Fit the model to `r_sp` in %.
     - `arch_model(100*r_sp, mean='Constant', vol='GARCH', p=1, q=1, dist='normal')`
  6. Compute the implied K.
  7. Draw a QQ plot of $\hat{z}_t=\hat{\varepsilon}_t/\hat{\sigma}_t$.
- **Interpretation**: what does the QQ plot say about the Normal distribution of z_t?
- **Report**: moment conditions, fitted parameters, implied kurtosis and a QQ plot.

A11 [Proposed]: The Hill Estimator as a Maximum Likelihood Estimator — Task

- **Model**: A5 [Solved]; notebook section A11
- **Assumption**: given the threshold $u = X_{(k+1)}$, the k exceedances are the order statistics of an i.i.d. Pareto sample
  - Pareto density $f(x) = \alpha u^{\alpha} x^{-\alpha-1}$, $x > u$, with tail index α
- **Concepts**
  - the log-likelihood $\ell(\alpha) = \sum_{i=1}^{k} \ln f(X_{(i)})$: how plausible the data are for a given value of α
  - the maximum likelihood estimator: the value of α that maximises $\ell(\alpha)$
  - ℓ', ℓ'': the first and second derivatives with respect to α
  - the Fisher information $-\mathbb{E}\,\ell''(\alpha)$; $\mathrm{SE}(\hat{\alpha}) \approx 1/\sqrt{\text{information}}$
- **Tasks**
  1. Write the log-likelihood $\ell(\alpha)$ of the k exceedances (the ordering adds only the constant $\ln k!$).
  2. Solve $\ell'(\alpha) = 0$ and show that the maximiser is the Hill estimator $\hat{\alpha} = k / \sum_{i=1}^{k} \ln(X_{(i)}/u)$ Hill (1975).
  3. Compute $\ell''(\alpha)$, the information k/α^2 and $\mathrm{SE}(\hat{\alpha}) \approx \hat{\alpha}/\sqrt{k}$.
  4. Check the SE on the data of A5.
  5. Simulate $n = 2\,000\,000$ values $|X|$, X Student-t with $\nu = 3$ (`numpy.random.default_rng(1)`).
  6. Plot $\hat{\alpha}_k$ (the Hill estimate from k values) against k/n and report $k = 200,\ 2\,000,\ 20\,000,\ 200\,000$.
- **Interpretation**: why does $\hat{\alpha}_k$ drift away from 3 when k is large?

Back Exercise list

B1(d) [Solved]: Quantile-Based Moments with Bootstrap Intervals (1/2)

- **Question**: do negative skewness and excess kurtosis survive measures that need no moments?
- **Measures** Kim & White (2004), both 0 for the Normal distribution (q_p: the quantile of order p)
  - Hinkley skewness $\frac{q_{0.95}+q_{0.05}-2q_{0.5}}{q_{0.95}-q_{0.05}}$: the position of the median between $q_{0.05}$ and $q_{0.95}$
  - it lies in $[-1, 1]$; negative: the left tail is longer than the right one
  - Crow–Siddiqui kurtosis K_{CS} (definition)
- **Block bootstrap**: 1 000 resamples (`default_rng(42)`)
  - draw $\lceil T/20 \rceil$ starting days uniformly ($\lceil \cdot \rceil$: rounding up) and take 20 consecutive returns from each
  - the blocks keep the volatility clustering inside them
  - cut to T and recompute both measures
  - the 95% percentile interval: the 2.5% and 97.5% quantiles of the 1 000 recomputed values

Back Exercise list

B1(d) [Solved]: Quantile-Based Moments with Bootstrap Intervals (2/2)

	Moment skew.	Hinkley [95% CI]	Excess kurt.	Crow–Siddiqui [95% CI]
BET-TR	−1.40	−0.036 [−0.085, 0.025]	17.8	1.06 [0.73, 1.43]
Bitcoin	−0.70	−0.018 [−0.069, 0.032]	12.0	2.25 [1.88, 2.64]
EUR/RON	0.73	0.041 [−0.022, 0.093]	15.1	4.87 [3.74, 6.13]

- In square brackets: 95% block-bootstrap confidence intervals (CI)
- Every kurtosis interval lies above 0: heavy tails are established without moments
- Moment skewness and Hinkley skewness measure different things
  - every Hinkley interval contains 0
  - this does not establish symmetry

MFM_ch1_inference

B7 [Proposed]: A Kurtosis That Does Not Converge — Task

- **Model**: B1 [Solved]; notebook section B7
- **Data**: `r_sp`; window n = the returns dated before 3 January of year $1990 + n$, $n = 1, \ldots, 36$
  - 253 returns for $n = 1$, 2 527 for $n = 10$, 9 067 for $n = 36$
- **Simulation**: 500 Student-t paths of length 9 245 for $\nu = 3$ and for $\nu = 6$ (`default_rng(42)`)
  - cut at the same window counts
- **Tasks**
  1. Compute the excess kurtosis of the S&P 500 on each window.
  2. List the three largest jumps, with the days that cause them.
  3. For each ν compute the 5%, 50% and 95% quantiles of the simulated kurtosis at each n.
  4. Plot the data curve over both bands on a symmetric log scale.
- **Interpretation**
  - why does the $\nu = 6$ band settle slowly although the kurtosis of t_6 is finite?
  - what does "excess kurtosis = 10.9" then mean?
- **Report**: expanding-window kurtosis, simulation quantile bands and the three largest jumps.

Back Exercise list

B8 [Proposed]: Stylised Facts Reproduced by GARCH-t — Task (1/2)

- **Model**: B3 and B6 [Solved], A10; notebook section B8
- **Data**: `r_sp` converted to % ($T = 9\,245$)
  - constant mean
  - ω in $\%^2$ a day

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

B8 [Proposed]: Stylised Facts Reproduced by GARCH-t — Task (2/2)

Tasks

1. Fit the model and list all five parameters with units.
2. Pilot: simulate one path and compute its excess kurtosis.
3. Then simulate 200 paths.
4. For each path compute these statistics.
   - the excess kurtosis
   - the ACF of $|r_t|$ at lags 1, 20, 100, 250
   - the Hill $\hat{\alpha}$ of $|r_t|$ at $k = 2\%$ of T
   - $L(1) = \mathrm{corr}(r_t, |r_{t+\mathbf{1}}|)$: the correlation between today's return and tomorrow's size (the leverage effect, B9)
5. Build the pointwise 5–95% intervals.
6. Mark each observed value as inside or outside.

Interpretation

- which facts are not reproduced?
- which model component would you change for each?

Report: observed statistics beside simulation intervals and proposed model changes.

Back Exercise list

B9 [Proposed]: Leverage Effect, BET-TR versus Bitcoin — Task

- **Model**: B3 [Solved]; notebook section B9
- **Lag convention**: $L(k) = \text{corr}(r_t, |r_{t+k}|)$ Christie (1982)
  - $L(k) < 0$: after falls, later returns are larger in absolute value (the leverage effect)
  - k counts observation days of each series
  - $k > 0$: today's return and a later size
  - example, $k = 1$: the BET-TR return of Monday 16 Mar 2020 is paired with $|r|$ of Tuesday 17 Mar 2020
- **Data**: `r_bet` and `r_btc`, each on its own calendar
- **Tasks**
  1. Compute $L(k)$ for $k = -20, \ldots, 20$.
  2. Plot both series with the i.i.d. reference bands $\pm 1.96/\sqrt{T}$ (0.036 and 0.030).
- **Tasks** (continued)
  3. Report $L(1)$ and $L(5)$ for both markets.
  4. State whether they lie outside the reference bands.
  5. Give two competing explanations for a weaker Bitcoin effect.
- **Interpretation**
  - why is a reference band not a confidence interval under clustering?
  - what does B10 add?

Back Exercise list

B10 [Proposed]: The Leverage Effect as a Regression — Task (1/2)

- **Model**: B4 [Solved], B9; notebook section B10
- **Regression**: $|r_{t+1}| = c + \beta^+ r_t^+ + \beta^- r_t^- + \sum_{j=1}^{5} \delta_j |r_{t-j}| + u_{t+1}$
  - $r_t^+ = \max(r_t, 0) \geq 0$, $r_t^- = \min(r_t, 0) \leq 0$: the positive and the negative part of the return
  - c: the constant; δ_j: the effect of $|r_{t-j}|$, controls for volatility clustering; u_{t+1}: the error
  - a gain $+a$ adds $\beta^+ a$, a loss $-a$ adds $-\beta^- a$
  - a symmetric response means $\beta^+ = -\beta^-$
  - design row for day t: $(1,\ r_t^+,\ r_t^-,\ |r_{t-1}|, \ldots, |r_{t-5}|)$
  - response: $|r_{t+1}|$
- **Data**: `r_sp`, `r_bet`, `r_btc`, `r_eur`

B10 [Proposed]: The Leverage Effect as a Regression — Task (2/2)

- **Tasks**
  1. Estimate by OLS with Newey–West standard errors Newey & West (1987), robust to heteroskedasticity and autocorrelation, with lag truncation $\lfloor 4(T/100)^{2/9} \rfloor$ (the number of error lags taken into account).
  2. Test $H_0 : \beta^+ + \beta^- = 0$ (a symmetric response) at 5%, against χ^2_1, with the Wald statistic
     - $W = (\hat{\beta}^+ + \hat{\beta}^-)^2 / [\mathrm{Var}(\hat{\beta}^+) + \mathrm{Var}(\hat{\beta}^-) + 2\,\mathrm{Cov}(\hat{\beta}^+, \hat{\beta}^-)]$
     - denominator: the variance of $\hat{\beta}^+ + \hat{\beta}^-$, from the Newey–West covariance matrix; W is the square of a t-statistic
  3. Plot the predicted $|r_{t+1}|$ after a gain $+a$ and a loss $-a$, lag controls at their means.
  4. Show why $|r_t|$ cannot be added as a control.
- **Report**: coefficient table, Wald test, response plot and collinearity explanation.

B11 [Proposed]: EUR/RON Across Regimes — Task (1/3)

- **Model**: B1, B4 [Solved]; notebook section B11
- **Data**: `r_eur` (BNR reference rate), split into three periods
  - Jul 2005 – 2011, 2012–2019, 2020–2026: 1 656, 2 013, 1 683 returns
  - lags built inside each period
- **Tasks**
  1. In each period regress $r_t = a + \rho\, r_{t-1} + e_t$ by OLS, with a heteroskedasticity- and autocorrelation-consistent (HAC) SE, as in B10.
     - a: the constant; ρ: the lag-1 autocorrelation; e_t: the error
     - lag truncation: $\lfloor 4(T/100)^{2/9} \rfloor$
  2. Report $\hat{\rho}$, the SE, the volatility and the excess kurtosis of each period.

B11 [Proposed]: EUR/RON Across Regimes — Task (2/3)

Tasks (continued)

3. Test equal ρ across periods with the statistic below.

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

The reference assumes asymptotically independent estimates across the three periods.

- $j = 1, 2, 3$: the period; weights $w_j = 1/\mathrm{SE}_j^2$, so precise estimates count more
- $\bar{\rho} = \sum_j w_j \hat{\rho}_j / \sum_j w_j$: the weighted mean of the three estimates
- W measures the spread of the estimates around $\bar{\rho}$; 2 degrees of freedom = 3 periods − 1

B11 [Proposed]: EUR/RON Across Regimes — Task (3/3)

- **Tasks** (continued)
  4. Separately, locate one variance break with the cumulative sum of squares (CUSUM) Inclán & Tiao (1994).
     - $C_k = \sum_{t \le k} x_t^2$: sum of squares up to day k; x_t is the return minus its full-sample mean
     - $D_k = C_k / C_T - k/T$: the share of total variance accumulated by day k, minus the share of time elapsed
     - $D_k \approx 0$ for every k if the variance is constant
     - the break is at $\arg\max_k |D_k|$, the day k where $|D_k|$ is largest
- **Interpretation**
  - why is the 2020–2026 SE so large?
  - why is the CUSUM date descriptive rather than a test?

B12 [Proposed]: Long Memory or Breaks? — Task (1/3)

- **Model**: B4, B6 [Solved]; notebook section B12 (helper `local_whittle(x, m)`)
- **Memory parameter** d: how fast the ACF of $x_t = |r_t|$ decays
  - $d = 0$: short memory, the ACF decays geometrically
  - $0 < d < 0.5$: stationary long memory, the ACF decays slowly, like k^{2d-1}
  - $d \geq 0.5$: the series is no longer stationary
- **Data**: $|r_t|$ of the S&P 500, BET-TR and Bitcoin

Back Exercise list

B12 [Proposed]: Long Memory or Breaks? — Task (2/3)

- **Local Whittle method** Robinson (1995): demean $x_t = |r_t|$
  - Fourier frequencies $\lambda_j = 2\pi j/T$, $j = 1, \ldots, m$; low frequencies correspond to long cycles
  - periodogram $I(\lambda) = \frac{1}{2\pi T}\left|\sum_t x_t e^{-\mathrm{i}\lambda t}\right|^2$: the strength of cycles of frequency λ in the series
  - long memory makes $I(\lambda_j)$ grow like λ_j^{-2d} at low frequencies
- $\hat{d}$ minimises over $d \in [-0.49, 0.99]$

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

  - $m = \lfloor T^p \rfloor$: the number of low frequencies used; $p \in (0,1)$ sets the bandwidth
  - SE $= 1/(2\sqrt{m})$

B12 [Proposed]: Long Memory or Breaks? — Task (3/3)

- **Tasks**
  1. Estimate d with $p = 0.65$ and give the 95% interval.
  2. Plot $\hat{d}$ against p from 0.45 to 0.80, with the lines $d = 0$ and $d = 0.5$.
  3. Locate one variance break of r_t (B11).
  4. Re-estimate $\hat{d}$ on each side of the break.
  5. Plot the ACF of $|r_t|$ before and after the break.
- **Interpretation**
  - is $\hat{d} > 0$ robust to p and to the break?
  - or is it consistent with regimes Diebold & Inoue (2001)?
  - what does $\hat{d} \geq 0.5$ imply?
- **Report**: memory-parameter estimates and intervals, bandwidth plot, break date and ACF before/after.

B13 [Proposed]: BET-TR vs S&P 500, Same Period — Task (1/2)

- **Model**: A1, B1 [Solved]; notebook section B13
- **Data**: BET-TR (total return, RON) and S&P 500 (price index, USD)
  - align the *prices* on common dates (23 Sep 2014 – 18 Sep 2026)
  - then compute 2 903 returns
  - interval lengths: 2 231 one-day, 538 of two or three days, 134 longer (at most 6 days)

Back Exercise list

B13 [Proposed]: BET-TR vs S&P 500, Same Period — Task (2/2)

- **Tasks**
  1. Plot the rebased wealth (log scale) and the two underwater curves, DD_t of A2(a).
  2. Compute these measures for both indices.
     - CAGR $= (P_{\mathbf{end}}/P_{\mathbf{start}})^{\mathbf{1}/Y} - 1$, with Y the number of calendar years; the volatility of simple returns ($A = 252$ by convention; the common days give about 242 returns a year)
     - the Sharpe ratio $252\bar{R}/(\sqrt{252}\,\hat{\sigma}_R)$ with $r_f = 0$: the annual return per unit of volatility
     - $\bar{R}$, $\hat{\sigma}_R$: the mean and the standard deviation of daily simple returns; r_f: the risk-free rate
     - the Sortino ratio Sortino & van der Meer (1991) $252\bar{R}/(\sqrt{252}\,\hat{\sigma}_D)$, with the downside deviation $\hat{\sigma}_D = \sqrt{\frac{\mathbf{1}}{T}\sum_t \min(R_t, 0)^{\mathbf{2}}}$ over all days
     - the Sortino ratio penalises loss days only
     - the MDD and the Calmar ratio (A2(a))
- **Report**: the chart and a two-column table
- **Interpretation**: list three reasons why the comparison is descriptive only

B14 [Proposed]: Significance of the Sharpe Difference — Task (1/2)

- **Model**: B1(d), B13; notebook section B14
- **Reference**: the standard error of the Sharpe ratio, Lo (2002)
- **Data**: the 2 903 paired returns of B13
  - simple returns for the Sharpe ratios (computed in B13) and the test correlation
  - log returns for the timing comparisons in task 5
- **Tasks**
  1. Convert to daily Sharpe ratios ($\widehat{SR}/\sqrt{252}$).
  2. Compute the Jobson–Korkie test Jobson & Korkie (1981) with the correction of Memmel (2003).
     - $z = (\widehat{SR}_1 - \widehat{SR}_2)/\sqrt{\hat{V}}$
     - $\hat{V} = \frac{1}{T}\left[2(1-\hat{\rho}) + \frac{1}{2}(\widehat{SR}_1^2 + \widehat{SR}_2^2 - 2\widehat{SR}_1\widehat{SR}_2\hat{\rho}^2)\right]$
     - $\widehat{SR}_1, \widehat{SR}_2$: the daily Sharpe ratios of BET-TR and the S&P 500; $\hat{\rho}$: the correlation of their returns
     - $\hat{V}$: the estimated variance of the difference; under H_0 (equal Sharpe ratios) $z \approx N(0,1)$

Back Exercise list

B14 [Proposed]: Significance of the Sharpe Difference — Task (2/2)

- **Tasks** (continued)
  3. Run a paired circular block bootstrap, 2 000 resamples (`default_rng(42)`).
     - draw $\lceil T/21 \rceil$ start days, take 21 consecutive days from each (wrapping at the end), cut to T
     - apply the **same** indices to both series
  4. Report the 95% percentile interval of the annualised difference.
  5. Compute four correlations.
     - same day
     - BET-TR with the S&P 500 of the previous day and of the next day
     - weekly (Friday) returns
- **Interpretation**
  - is the difference significant?
  - what do both methods assume?
  - which correlation belongs in a portfolio?

B15 [Proposed]: Range-Based Estimators on the S&P 500 — Task (1/2)

- **Model**: A8 [Solved]; notebook section B15
  - tested helpers for all five estimators
  - formulas in Chapter 1
- **Data**: S&P 500 index OHLC, 2 Jan 2008 – 18 Sep 2026
- SPY ETF on one adjusted basis
  - C^* = the adjusted close (for dividends and splits)
  - $O^*, H^*, L^* = (C^*/C) \times O, H, L$, with C the unadjusted close and C^*/C the day's adjustment factor
  - example, 16 Mar 2020: factor $218.08/239.85 = 0.9092$; $O^* = 241.18 \times 0.9092 = 219.29$, $H^* = 233.58$, $L^* = 215.82$

B15 [Proposed]: Range-Based Estimators on the S&P 500 — Task (2/2)

- **Tasks**
  1. Compute 21-day rolling close-to-close, Parkinson, Garman–Klass Garman & Klass (1980), Rogers–Satchell Rogers & Satchell (1991) and Yang–Zhang Yang & Zhang (2000) volatilities ($A = 252$).
  2. Compute their time averages.
  3. For the index and for SPY, compare the 21-day averages of close-to-close and Yang–Zhang with two full-sample volatilities.
     - overnight: $\mathrm{sd}(o_t - c_{t-\mathbf{1}})\sqrt{252}$, with $o_t - c_{t-\mathbf{1}}$ the overnight log return (notation of A8)
     - open-to-close: $\mathrm{sd}(c_t - o_t)\sqrt{252}$
     - sd: the sample standard deviation
- **Interpretation**: why does Yang–Zhang stay below close-to-close on the index but not on SPY?
- **Report**: a table of average volatilities and an explanation of the overnight contribution.

C2 [Proposed]: Did BET-TR Autocorrelation Fall after the FTSE Reclassification? (1/3)

- **Model**: B4 [Solved], B11; a project seed
- **Context**: on less liquid markets, index returns often have a positive lag-1 autocorrelation, because the prices of the constituents are stale
- **Hypothesis to test**: after the reclassification, liquidity rose and the BET-TR lag-1 autocorrelation fell
  - Romania became a Secondary Emerging market of FTSE (Financial Times Stock Exchange, FTSE Russell, index provider) Russell, an international index provider, on 21 Sep 2020
- **Data**: `r_bet` and a liquidity measure
  - liquidity: daily traded value (volume $\times$ raw close, in lei) of 11 BVB blue chips
  - the stocks, by BVB ticker: TLV (Banca Transilvania), SNP (OMV Petrom), BRD (BRD–Groupe Société Générale), TGN (Transgaz), FP, SNG (Romgaz), H2O (Hidroelectrica), SNN (Nuclearelectrica), EL (Electrica), DIGI, TEL (Transelectrica); files `<ticker>.RO.csv`, fields `volume` and `close`
  - summed over the stocks listed that day
  - report how many are listed each year

Back Exercise list

C2 [Proposed]: Did BET-TR Autocorrelation Fall after the FTSE Reclassification? (2/3)

Tasks

1. Compute $\hat{\rho}_1$ on rolling 250-day windows (step 5 days).
2. Add the 5% test band $\pm 1.96\,\mathrm{SE}$ around zero, with the robust SE of each window computed as in B4 under $\rho_1 = 0$; it tests $H_0 : \rho_1 = 0$ and is not a confidence interval around $\hat{\rho}_1$.
3. Estimate ρ_1 before and after 21 Sep 2020 as an AR(1) slope with HAC SEs (B11).
4. Report the statistic for the difference of the two estimates.

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

- $\mathrm{SE}_1, \mathrm{SE}_2$: the HAC standard errors before and after
- under H_0 (no change) $z \approx N(0, 1)$, if the two periods are independent

C2 [Proposed]: Did BET-TR Autocorrelation Fall after the FTSE Reclassification? (3/3)

- **Tasks** (continued)
  5. Report the interaction e in the regression below (HAC, 10 lags).

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

  - $\text{post}_t = 1$ after 21 Sep 2020 and 0 before
  - b: the autocorrelation before; e: its change after the reclassification; a, c: the constant and its change

  6. Correlate the rolling $\hat{\rho}_1$ with the log 250-day average traded value.

- **Deliverable**: a Quantlet folder and one slide
  - a descriptive conclusion and its main limitation
  - not a causal claim

C3 [Proposed]: Is the VIX Mostly Path-Dependent? — Task (1/5)

- **Model**: the lecture case study
- **Question**: can past S&P 500 returns alone explain the VIX, as Guyon & Lekeufack (2023) find?
  - 94.6% of its variance in 2000–2018, 85.5% in 2019 – May 2022
- **Data**: daily closes of the S&P 500, the VIX and the 9-day VIX
  - S_{t_i}: the S&P 500 level at the close of day t_i; simple returns $r_{t_i} = S_{t_i}/S_{t_{i-1}} - 1$
  - volatility indices divided by 100
- **Model** (Guyon & Lekeufack (2023), eq. (3.7)): today's volatility is a linear function of two summaries of past returns

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

  - σ_t: the VIX (or the 9-day VIX) at the close of day t, divided by 100
  - $\beta_0, \beta_1, \beta_2$: coefficients estimated from the data

Back Exercise list

C3 [Proposed]: Is the VIX Mostly Path-Dependent? — Task (2/5)

- **Trend feature**: a weighted sum of the returns of the last 1 000 days

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

  - i: the age of the return, in days; $\Delta t = 1/252$: one day, in years; $K_1(i\Delta t)$: its weight
  - $R_{1,t} < 0$ after a run of falls, $R_{1,t} > 0$ after a run of rises
- **Volatility feature**: the square root of a weighted sum of squared returns

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

  - a weighted historical volatility, with weights K_2
- **Reading the coefficients**
  - $\beta_1 < 0$: after falls ($R_{1,t} < 0$) volatility rises, i.e. the leverage effect
  - $\beta_2 > 0$: high recent volatility raises today's volatility, i.e. volatility clustering

C3 [Proposed]: Is the VIX Mostly Path-Dependent? — Task (3/5)

- **Kernels**, eq. (3.9) in the paper: a time-shifted power law

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

  - τ: the age, in years; Z_j such that $\sum_{i=0}^{999} K_j(i\Delta t)\Delta t = 1$
  - $\alpha_j > 1$: the speed at which the weights decay (not the tail index); a small α_j means slowly decaying weights over the 1 000-day window
  - $\delta_j > 0$: the shift; a small δ_j puts the weight on the last few days
- **Parameters and units**
  - β_0 in volatility units ($0.03 = 3$ VIX points)
  - β_1, β_2 unitless
  - δ_j in years
- **Start** (Table 3 in the paper, VIX row)
  - $\beta_0 = 0.057$, $\beta_1 = -0.095$, $\beta_2 = 0.82$
  - $\alpha_1 = 1.06$, $\delta_1 = 0.020$, $\alpha_2 = 1.60$, $\delta_2 = 0.052$
  - bounds $\alpha_j \in [1.0001, 20]$, $\delta_j \in [10^{-4}, 5]$

C3 [Proposed]: Is the VIX Mostly Path-Dependent? — Task (4/5)

- **Tasks**, in stages
  1. Fixed kernels: compute $R_{\mathbf{1},t}$ and Σ_t with the Table 3 kernels.
  2. Estimate $\beta_{\mathbf{0}}, \beta_{\mathbf{1}}, \beta_{\mathbf{2}}$ by OLS on 2000–2018.
  3. Full model: estimate all seven parameters by nonlinear least squares on 1 Jan 2000 – 31 Dec 2018, minimising $\sum_t (\sigma_t - \hat{\sigma}_t)^2$, with $\hat{\sigma}_t$ the model value.
     - use `scipy.optimize.least_squares`
     - recompute the features at every trial value
  4. Report R^2 and the root mean squared error (RMSE, defined on the next slide) on the training set and on 1 Jan 2019 – 15 May 2022.
     - for the full model and for the restrictions $\beta_{\mathbf{1}} = 0$ and $\beta_{\mathbf{2}} = 0$; re-estimate the remaining parameters by nonlinear least squares; the dropped feature's kernel is irrelevant
  5. Repeat the full model for the 9-day VIX (training from 2011).
  6. Keep the parameters fixed and score 16 May 2022 – 18 Sep 2026.

C3 [Proposed]: Is the VIX Mostly Path-Dependent? — Task (5/5)

- **Scores**, computed separately on each sample of T days
  - RMSE $= \sqrt{\frac{1}{T}\sum_t(\sigma_t - \hat{\sigma}_t)^2}$: the typical error, in volatility units
  - $R^2 = 1 - \sum_t(\sigma_t - \hat{\sigma}_t)^2 / \sum_t(\sigma_t - \bar{\sigma})^2$: the share of the variation of σ_t explained by the model
  - $\bar{\sigma}$: the mean of σ_t; $R^2 = 1$: a perfect fit; $R^2 \leq 0$: no better than the mean
- **Report**
  - the parameter table and the scores next to Table 2 of the paper
  - the observed and fitted VIX, and the ratio of observed to fitted VIX
  - the fitted kernels
- **Interpretation**: the model fits the VIX on the same day
  - it is not a next-day forecast
  - which feature explains more, $R_{1,t}$ or Σ_t?

C3: What Does Weight Normalisation Change?

- The weights summing to one are $w_{j,i} = K_j(i\Delta t)\Delta t$, with $\Delta t = 1/252$.
- The formulas of task 1 use K_j without the factor Δt: the trend feature is 252 times the average computed with weights $w_{1,i}$.
- The volatility feature is $\sqrt{252}$ times the square root of the weighted mean of squared returns, computed with $w_{2,i}$.
- Example: if every return is 0.1%, the weighted mean is 0.001, but the trend feature is 0.252.
- Model coefficients depend on this scale: use the same normalisation for estimation and over the test period.

References (1/3)

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

References (2/3)

- Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357–384.
- Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 3(5), 1163–1174.
- Inclán, C., Tiao, G.C. (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association*, 89(427), 913–923.
- Jarque, C.M., Bera, A.K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3), 255–259.
- Jobson, J.D., Korkie, B.M. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance*, 36(4), 889–908.
- Kim, T.-H., White, H. (2004). On more robust estimation of skewness and kurtosis. *Finance Research Letters*, 1(1), 56–73.
- Ledoit, O., Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, 15(5), 850–859.
- Ljung, G.M., Box, G.E.P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303.
- Lo, A.W. (2002). The statistics of Sharpe ratios. *Financial Analysts Journal*, 58(4), 36–52.
- Lobato, I., Nankervis, J.C., Savin, N.E. (2001). Testing for autocorrelation using a modified Box–Pierce Q test. *International Economic Review*, 42(1), 187–205.
- 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.

References (3/3)

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

Acronyms used in this material (1/2)

- **ACF**: AutoCorrelation Function
- **AI**: Artificial Intelligence
- **AR**: AutoRegressive model
- **ARCH-LM**: ARCH Lagrange Multiplier (test)
- **BET-TR**: BET Total Return (index)
- **BNR**: Banca Națională a României — *the National Bank of Romania*
- **BVB**: Bursa de Valori București — *the Bucharest Stock Exchange*
- **CAGR**: Compound Annual Growth Rate
- **CI**: Confidence Interval
- **COVID-19**: COronaVIrus Disease 2019
- **CUSUM**: CUmulative SUM (of squares)
- **EGARCH**: Exponential GARCH
- **ETF**: Exchange-Traded Fund
- **FIGARCH**: Fractionally Integrated GARCH
- **FTSE**: Financial Times Stock Exchange (FTSE Russell, index provider)

Acronyms used in this material (2/2)

- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity
- **GJR-GARCH**: Glosten–Jagannathan–Runkle GARCH
- **HAC**: Heteroskedasticity and Autocorrelation Consistent (standard errors)
- **HAR**: Heterogeneous AutoRegressive (model)
- **JB**: Jarque–Bera (test)
- **MDD**: Maximum DrawDown
- **OHLC**: Open, High, Low, Close
- **OLS**: Ordinary Least Squares
- **QQ**: Quantile–Quantile (plot)
- **RMSE**: Root Mean Squared Error
- **SE**: Standard Error
- **US**: United States
- **UTC**: Coordinated Universal Time (Temps Universel Coordonné)
- **VIX**: Cboe Volatility Index