

Statistics of Financial Markets

Chapter 10: VaR, ES and backtesting



Daniel Traian PELE

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

Acronyms used in this material

- **AI**: Artificial Intelligence
- **BCBS**: Basel Committee on Banking Supervision
- **BET**: Bucharest Exchange Trading (index)
- **BET-TR**: BET Total Return (index)
- **BIS**: Bank for International Settlements
- **BVB**: Bursa de Valori București (the Bucharest Stock Exchange)
- **CLT**: Central Limit Theorem
- **EODHD**: EOD Historical Data (market-data provider)
- **ES**: Expected Shortfall
- **EVT**: Extreme Value Theory
- **EWMA**: Exponentially Weighted Moving Average
- **FHS**: Filtered Historical Simulation
- **FRTB**: Fundamental Review of the Trading Book
- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity
- **GPD**: Generalised Pareto Distribution
- **HS**: Historical Simulation
- **LR**: Likelihood Ratio (test)
- **MAR32**: Market Risk standard, chapter 32 of the Basel Framework (backtesting)
- **MAR33**: Market Risk standard, chapter 33 of the Basel Framework (capital calculation)
- **POF**: Proportion Of Failures (Kupiec test)
- **POT**: Peaks Over Threshold
- **SFE**: Statistics of Financial Markets, Exercises (Quantlet collection of the textbook)
- **US**: United States

Today's Question and Route

- **Question**: how much can a portfolio lose tomorrow, and how do we know that the answer was right?
  - Chapter 5 described the tails, Chapter 9 the volatility of tomorrow; this chapter turns both into risk numbers and tests them
- **Route** of the chapter
  - VaR (value at risk) and ES (expected shortfall): definitions, coherence, elicitability
  - estimation: historical simulation, Normal, Student-t, Cornish–Fisher, EVT, GARCH and filtered historical simulation
  - portfolios: variance–covariance, tail dependence, a first look at copulas
  - backtesting: Kupiec, Christoffersen, the Basel traffic light, ES backtesting

Learning Outcomes

- Define VaR and ES at a tail probability α (VaR 1%, ES 2.5%) and compute them for the Normal and Student-t distributions
- State the axioms of a coherent risk measure and show with a counterexample that VaR is not subadditive
- Estimate VaR and ES by historical simulation, parametric methods, Cornish–Fisher, EVT, GARCH and filtered historical simulation
- Compute a portfolio VaR and explain why correlation does not measure the risk of joint crashes
- Backtest a VaR model with the Kupiec and Christoffersen tests and the Basel traffic light

Reading and Tools

- Textbook: Franke, Härdle & Hafner (2019), Ch. 16 (VaR and backtesting), Ch. 17 (copulae and VaR), Sections 18.1 and 18.4 (extreme risks)
  - exercises with solutions: Borak, Härdle & López-Cabrera (2013), Ch. 16
  - risk measures and copulas: McNeil, Frey & Embrechts (2015), Ch. 2 and 7
- Python Quantlets of this chapter: Quantlets/Ch_10
  - ported from the SFE Quantlets VaRest, SFEVaRbank, SFEVaRtimeplot, SFEVaRqqplot and SFEvar_pot_backtesting
- Lecture notebook: open in Google Colab
- Video course: Measuring Statistical Risk

The 4:15 Report and RiskMetrics

- Around 1990: J.P. Morgan's chairman, Dennis Weatherstone, asks for one number, every day at 4:15 pm, summarising how much the bank could lose the next day
  - the answer became the **value at risk** (VaR): a loss that is exceeded only with a small probability
- October 1994: the methods and data are published free of charge as RiskMetrics J.P. Morgan/Reuters (1996)
  - Normal returns, EWMA volatility (Chapter 8), VaR at a 5% tail probability over one day
  - VaR became the common language of banks, regulators and fund managers



The J.P. Morgan building at 23 Wall Street, New York

Photo: Beyond My Ken (2012); CC BY-SA 4.0; Wikimedia Commons

RiskMetrics and the Basel Rules

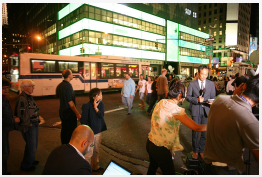
- The **BCBS** (Basel Committee on Banking Supervision) sits at the **BIS** (Bank for International Settlements) in Basel history of the Basel Committee
  - 1996: banks may compute market-risk capital with their own VaR models BCBS (1996a)
  - rule: VaR 1% over 10 days, times a multiplier of at least 3; the multiplier grows when the model fails the backtest BCBS (1996b)
- 2019: the **FRTB** (Fundamental Review of the Trading Book) BCBS (2019)
  - capital from ES 2.5% instead of VaR 1% (MAR33, BCBS (2023b)); backtesting still counts VaR exceptions (MAR32, BCBS (2023a))



The BIS tower in Basel, seat of the Basel Committee

Photo: Taxiarchos228 (2014); CC BY-SA 4.0; Wikimedia Commons

The Crises of 2008 and 2020



Lehman Brothers headquarters, New York, 15 September 2008

Photo: Robert Scoble (2008); CC BY 2.0; Wikimedia Commons



The Financial District of New York, 25 March 2020

Photo: Billie Grace Ward (2020); CC BY 2.0; Wikimedia Commons

- The worst day of the S&P 500 since 2000: -12.8% on 16 March 2020; a Normal model with the sample volatility gives such a day a probability of practically zero
- A risk measure must be judged in such periods: we use 2008 and 2020 as stress tests throughout the chapter

Losses, Horizon and Level

- X: the **P&L** (profit and loss) or the return of a position over a horizon h (here one day, in %)
  - the loss is $L = -X$: a positive number when the position loses money
- The **level** α is the probability of the tail: $\alpha = 1\%$ or $\alpha = 2.5\%$
  - we write **VaR 1%** and **ES 2.5%**: the number in the name is always the tail probability
- $q_\alpha(X) = \inf\{x : P(X \leq x) > \alpha\}$: the α-**quantile** of X (Chapter 2)
  - for a continuous distribution function F: $q_\alpha = F^{-1}(\alpha)$, a negative number for small α

Value at Risk

- **Definition**: $\mathrm{VaR}_\alpha(X) = -q_\alpha(X)$
  - the loss that is exceeded with probability α: $P(L > \mathrm{VaR}_\alpha) \leq \alpha$
  - the minus sign makes VaR a positive loss; a VaR of 3% means "a loss of more than 3% happens on 1 day in 100"
- What VaR does **not** say
  - how large the loss is on the bad days: a loss of 4% and a loss of 40% beyond VaR give the same VaR
  - VaR is not the maximum loss: by construction it is exceeded on about α of the days
- In money: a position of value W has $\mathrm{VaR} = W \times \mathrm{VaR}(\%)/100$

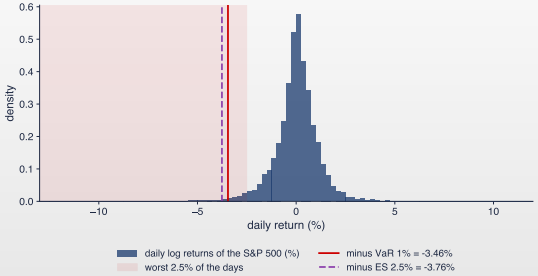
Expected Shortfall

- **Definition**: $\mathrm{ES}_\alpha(X) = -\frac{1}{\alpha}\int_0^\alpha q_u(X)\,du$: the average of the VaRs at all levels below α
  - for a continuous distribution: $\mathrm{ES}_\alpha = -E[X \mid X \le q_\alpha(X)]$, the average loss on the worst α of the days
  - u: the integration variable, a tail level running from 0 to α
  - other names: CVaR (conditional VaR), TVaR (tail VaR), expected tail loss
- Properties
  - $\mathrm{ES}_\alpha \ge \mathrm{VaR}_\alpha$ at the same level α
  - ES looks inside the tail: two distributions with the same VaR but different tails have different ES

Worked Example: VaR and ES from 100 Days

- A position has 100 daily returns; the five worst are -6.1, -4.8, -4.2, -3.9, -3.5 (in %)
  - empirical quantile at level α: the k-th smallest return, $k = \lceil n\alpha \rceil$ ($\lceil \cdot \rceil$: rounding up)
- **VaR 1%**: $k = \lceil 100 \times 0.01 \rceil = 1$, so $\mathrm{VaR}_{1\%} = 6.1\%$
  - one observation decides VaR 1%: with 100 days the estimate is very imprecise
- **ES 2.5%**: $k = \lceil 2.5 \rceil = 3$; the average of the three worst losses
  - $\mathrm{ES}_{2.5\%} = (6.1 + 4.8 + 4.2)/3 = 5.03\%$; $\mathrm{VaR}_{2.5\%} = 4.2\%$
- A position of 1 million RON: VaR 1% = 61 000 RON; on one day in a hundred the loss is larger

VaR 1% and ES 2.5% of the S&P 500



- Daily log returns since 2000 (6 714 days); historical estimates: VaR 1% = 3.46%, VaR 2.5% = 2.51%, ES 2.5% = 3.76%
- Interpretation: ES 2.5% averages the whole shaded tail, including the extreme days of 2008 and 2020; VaR 1% is one point in that tail
- For a position of 1 million RON in the index: VaR 1% about 34 552 RON, ES 2.5% about 37 613 RON

SFM_ch10_var_methods

The Normal Distribution: Closed Forms

- $X \sim N(\mu, \sigma^2)$, $z_\alpha = \Phi^{-1}(\alpha)$ (Φ, φ: distribution function and density of $N(0,1)$)
  - $\mathrm{VaR}_\alpha = -(\mu + \sigma z_\alpha), \quad \mathrm{ES}_\alpha = -\mu + \sigma \dfrac{\varphi(z_\alpha)}{\alpha}$
  - derivation: $E[Z \mid Z \le z] = -\varphi(z)/\Phi(z)$ because $\varphi'(z) = -z\varphi(z)$
- Numbers: $z_{0.01} = -2.326$; $z_{0.025} = -1.960$, $\varphi(z_{0.025}) = 0.0584$, $\varphi(z_{0.025})/0.025 = 2.338$
  - ES 2.5% / VaR 1% $= 2.338/2.326 = 1.005$ (with $\mu = 0$): for Normal returns the two measures are almost equal
  - this is why ES 2.5% replaced VaR 1% in Basel: about the same size for Normal data, larger for heavy tails

Worked Example: Normal VaR and ES of the S&P 500

- Sample moments of the daily returns since 2000: $\hat{\mu} = 0.025$, $\hat{\sigma} = 1.21$ (in %)
  - VaR 1% $= -(0.025 - 2.326 \times 1.21) = -0.025 + 2.822 = 2.80\%$
  - ES 2.5% $= -0.025 + 2.338 \times 1.21 = -0.025 + 2.836 = 2.81\%$
- Compare with the data (historical simulation): VaR 1% $= 3.46\%$, ES 2.5% $= 3.76\%$
  - the Normal VaR 1% is 19% too low: the Normal tail is too thin (Chapter 2)
  - the Normal ES 2.5% is even further from the data: it ignores the extreme days
- Interpretation: two parameters cannot describe the tail; the next methods change either the distribution or the data used

The Student-t Distribution: Closed Forms

- $X = m + s\,T$, $T \sim t_\nu$ (location m, scale s, ν degrees of freedom, Chapter 2); $t_\alpha = t_\nu^{-1}(\alpha)$, f_ν the density of t_ν
  - $\mathrm{VaR}_\alpha = -(m + s\,t_\alpha), \quad \mathrm{ES}_\alpha = -m + s\,\dfrac{f_\nu(t_\alpha)}{\alpha}\,\dfrac{\nu + t_\alpha^2}{\nu - 1}$ (for $\nu > 1$)
  - the variance is $s^2\nu/(\nu-2)$: a **standardised** t (variance 1) has $s = \sqrt{(\nu-2)/\nu}$
- **Worked example**: S&P 500, maximum likelihood: $\hat{m} = 0.068$, $\hat{s} = 0.703$, $\hat{\nu} = 2.67$
  - $t_{0.01} = -5.023$: VaR 1% $= -(0.068 + 0.703 \times (-5.023)) = 3.46\%$
  - ES factor 5.720: ES 2.5% $= -0.068 + 0.703 \times 5.720 = 3.95\%$
- Interpretation: with $\nu < 3$ the fitted tail is very heavy: VaR 1% matches the data, ES 2.5% is above the historical 3.76%

Scaling VaR from One to Ten Days

- **Square-root-of-time rule**: $\mathrm{VaR}^{(h)} \approx \sqrt{h}\,\mathrm{VaR}^{(\mathbf{1})}$
  - exact for i.i.d. Normal returns with mean 0: the h-day return has standard deviation $\sigma\sqrt{h}$
  - Basel 1996: 10-day VaR, usually scaled from one day with $\sqrt{10} = 3.162$
- S&P 500, historical simulation
  - one day: 3.46%; scaled: $3.162 \times 3.46 = 10.93\%$
  - VaR 1% of overlapping 10-day returns: 10.37%
- Why the rule fails in general
  - heavy tails: a sum of 10 heavy-tailed returns is closer to Normal (CLT, central limit theorem), so the ratio of quantiles is below $\sqrt{10}$
  - autocorrelation: $\mathrm{Var}(r_{t+\mathbf{1}} + \cdots + r_{t+\mathbf{10}}) = \sigma^{\mathbf{2}}\big(10 + 2\sum_{k=\mathbf{1}}^{\mathbf{9}}(10-k)\rho_k\big)$; S&P 500: $\hat{\rho}_{\mathbf{1}} = -0.10$, ratio 3.00; BET: $\hat{\rho}_{\mathbf{1}} = +0.11$, ratio $3.86 > \sqrt{10}$
  - volatility clustering: after a calm day the 10-day risk is larger than $\sqrt{10}$ times today's, after a storm smaller (Chapter 9)

1999: Four Axioms for a Risk Measure

- Artzner et al. (1999): a risk measure ρ maps a loss L to the capital that makes it acceptable; ρ is **coherent** if
  - **monotonicity**: $L_1 \leq L_2$ always $\Rightarrow \rho(L_1) \leq \rho(L_2)$
  - **translation invariance**: $\rho(L + c) = \rho(L) + c$ for a sure amount c
  - **positive homogeneity**: $\rho(\lambda L) = \lambda \rho(L)$, $\lambda > 0$
  - **subadditivity**: $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$
- Subadditivity: diversification never adds risk; a bank can add the capital of its desks as an upper bound



Freddy Delbaen, co-author of the axioms (1997)

Photo: ETH-Bibliothek Zürich, Bildarchiv (1997); CC BY-SA 4.0; Wikimedia Commons

Counterexample: Two Bonds, Step by Step

- Two independent bonds; each loses 100 with probability 4% (default) and 0 otherwise; level $\alpha = 5\%$
  - one bond: $P(L > 0) = 4\% \leq 5\%$, so $\mathrm{VaR}_{5\%} = 0$ for each
- Both bonds: $P(L = 200) = 0.16\%$, $P(L = 100) = 7.68\%$, so $P(L > 0) = 7.84\% > 5\%$
  - $\mathrm{VaR}_{5\%}(L_1 + L_2) = 100 > 0 + 0$: VaR **punishes** diversification
- ES 5% (the average of the worst 5% of outcomes)
  - one bond: $(4\% \times 100 + 1\% \times 0)/5\% = 80$
  - both: $(0.16\% \times 200 + (5 - 0.16)\% \times 100)/5\% = 103.2 \leq 80 + 80$: subadditive

Coherence of ES and VaR

- **ES** satisfies all four axioms Acerbi & Tasche (2002)
  - reason: ES is an average of quantiles over the tail, and averaging respects sums
- **VaR** satisfies the first three; subadditivity can fail
  - fails for discrete losses (credit, the bonds above) and for extremely heavy tails (tail index below 1, Chapter 5)
  - holds for elliptical distributions (Normal, Student-t) and, in practice, for most equity returns Daníelsson et al. (2013)
- ES needs a finite mean of the losses; VaR always exists

Elicitability and Forecast Evaluation

- A statistic is **elicitable** if some loss (scoring) function is minimised, on average, by its true value Gneiting (2011)
  - the mean: squared error; the median: absolute error
- **VaR is elicitable**: the quantile (pinball) loss $S(v, x) = (\alpha - 1\{x < -v\})(x + v)$
  - v: the VaR forecast; x: the realised return; $1\{\cdot\}$: 1 if the event occurs, 0 otherwise
  - two VaR models can be compared by their average quantile loss
- **ES alone is not elicitable** Gneiting (2011); the pair (VaR, ES) is Fissler & Ziegel (2016)
  - consequence: ES forecasts are tested together with VaR forecasts (Section 7)

Summary of the Five Unconditional Methods

Method	Assumption	Weak point
historical simulation (HS)	the past window is a sample of tomorrow	few tail points; nothing beyond the worst day
Normal	returns are Normal	tails far too thin
Student-t	returns are Student-t	one tail shape for both tails
Cornish–Fisher	Normal quantile corrected by skewness and kurtosis	breaks down for large kurtosis
EVT (POT)	GPD above a high threshold (Chapter 5)	choice of the threshold

- "Unconditional": the same VaR every day of the window; conditional methods (Section 5) use today's volatility
- All methods: daily log returns in % (Chapter 1), VaR and ES as positive losses

Historical Simulation

- **Algorithm**: take the last n returns $r_{t-n+\mathbf{1}}, \dots, r_t$, sort them: $r_{(\mathbf{1})} \le \cdots \le r_{(n)}$
  - $\widehat{\mathrm{VaR}}_\alpha = -r_{(k)}$, $k = \lceil n\alpha \rceil$; $\quad \widehat{\mathrm{ES}}_\alpha = -\frac{\mathbf{1}}{k} \sum_{i=\mathbf{1}}^{k} r_{(i)}$
  - no distribution is assumed: the empirical distribution of the window is the forecast
- **Worked example**: S&P 500 since 2000, $n = 6\,714$
  - VaR 1%: $k = \lceil 67.14 \rceil = 68$, the 68-th worst day: 3.46%
  - ES 2.5%: the average of the $\lceil 167.85 \rceil = 168$ worst days: 3.76%
- The window n is a trade-off
  - short (250 days): reacts, but VaR 1% rests on 3 observations; long (all data): stable, but blind to today's volatility
  - **ghost effect**: a crash day changes VaR when it enters the window and again, abruptly, n days later when it leaves

Precision of a Historical VaR

- A VaR estimate is a statistic: it has a sampling error; we measure it with the bootstrap (Chapter 2)
  - resample the n returns with replacement 2000 times, recompute VaR and ES, read the 2.5% and 97.5% quantiles
- S&P 500, all 6 714 days: VaR 1% $= 3.46\%$, 95% interval [3.18, 3.66]; ES 2.5% $= 3.76\%$, [3.46, 4.07]
  - last 500 days: VaR 1% $= 2.75\%$, [1.99, 4.96]; ES 2.5% $= 2.94\%$, [2.14, 3.82]
- With two years of data the interval of VaR 1% is wider than 2 percentage points: VaR 1% rests on 5 observations
- Interpretation: differences between methods smaller than the bootstrap interval are not evidence that one method is better

SFM_ch10_var_methods

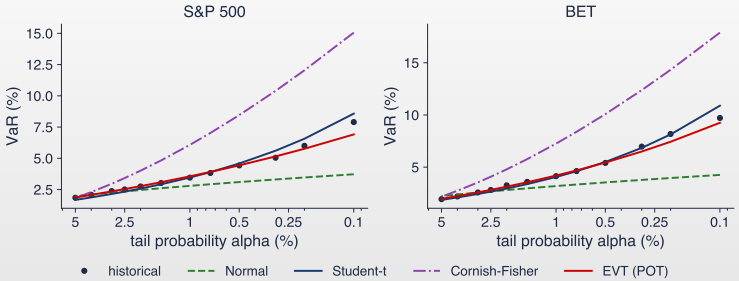
The Cornish–Fisher Expansion

- Cornish & Fisher (1938): correct the Normal quantile $z = z_\alpha$ with the skewness S and the excess kurtosis K (kurtosis minus 3)
  - $z_{CF} = z + \frac{(z^2-1)S}{6} + \frac{(z^3-3z)K}{24} - \frac{(2z^3-5z)S^2}{36}, \quad \mathrm{VaR}_\alpha = -(\mu + \sigma z_{CF})$
  - negative S and positive K push the quantile further into the left tail
- S&P 500: $S = -0.35$, $K = 10.7$, $z_{0.01} = -2.326$
  - $z_{CF} = -5.03$, so VaR 1% = 6.08%, against the historical 3.46%
- Interpretation: the expansion is a small correction around the Normal; with an excess kurtosis above 10 it overshoots badly
  - use it only when $|S|$ and K are moderate (for example monthly returns or diversified portfolios)

EVT: Peaks over Threshold (Chapter 5)

- Losses $L = -r$; threshold u = the 90% quantile of the losses McNeil & Frey (2000); excesses $L - u$ above u follow a GPD (generalised Pareto distribution) with shape ξ and scale β
  - $\mathrm{VaR}_\alpha = u + \frac{\beta}{\xi}\left[\left(\frac{n\alpha}{N_u}\right)^{-\xi} - 1\right]$, $\mathrm{ES}_\alpha = \frac{\mathrm{VaR}_\alpha + \beta - \xi u}{1 - \xi}$ (N_u: number of excesses)
  - n: the number of days; $n\alpha/N_u$: the tail probability relative to the share of excesses; $\xi > 0$: a heavy tail
- S&P 500: $u = 1.26\%$, $N_u = 672$ (10.0% of the days), $\hat{\xi} = 0.164$, $\hat{\beta} = 0.820$
  - VaR 1%: $n\alpha/N_u = 0.0999$, $0.0999^{-\hat{\xi}} = 1.461$; $\hat{\beta}/\hat{\xi} = 4.988$; $1.26 + 4.988 \times (1.461 - 1) = 3.56\%$
  - ES 2.5% = 3.77%; $\hat{\xi} > 0$: a heavy (Pareto-type) tail
- EVT smooths the tail and extrapolates beyond the worst observed day: useful for VaR 0.1% and for stress tests

VaR across Tail Probabilities



- Each line: VaR_α of one method for α from 5% to 0.1%; points: historical simulation
- Interpretation: at 5% all methods agree; at 1% the Normal is too low; at 0.1% Student-t and EVT stay close to the data, while Cornish–Fisher explodes

SFM_ch10_var_methods

VaR 1% by Method on Six Series

	HS	Normal	Student-t	Cornish–Fisher	EVT	K	$\hat{\xi}$
S&P 500	3.46	2.80	3.46	6.08	3.56	10.7	0.16
DAX	4.20	3.26	3.96	5.42	4.05	6.1	0.12
BET	4.14	3.19	4.06	7.25	4.19	11.1	0.23
Bitcoin	10.54	7.99	11.06	18.89	10.38	12.0	0.12
TLV	5.11	4.01	4.63	10.19	4.90	13.5	0.26
SNP	4.29	3.70	4.26	7.55	4.38	8.9	0.22

- Whole samples (S&P 500, DAX, BET since 2000; Bitcoin since 2014; TLV and SNP since 2010), in %; K: excess kurtosis
- HS, Student-t and EVT differ by at most 0.7 percentage points; the Normal VaR is 14–24% too low everywhere
- Cornish–Fisher gives 1.3 to 2.0 times the historical VaR: with K between 6 and 14 the expansion is outside its range

SFM_ch10_var_methods

ES 2.5% by Method on Six Series

	HS	Normal	Student-t	Cornish–Fisher	EVT	K	$\hat{\xi}$
S&P 500	3.76	2.81	3.95	6.65	3.77	10.7	0.16
DAX	4.26	3.27	4.37	5.81	4.24	6.1	0.12
BET	4.57	3.21	4.76	7.93	4.57	11.1	0.23
Bitcoin	10.84	8.03	13.37	20.69	10.90	12.0	0.12
TLV	5.37	4.03	5.03	11.20	5.35	13.5	0.26
SNP	4.70	3.72	4.57	8.18	4.71	8.9	0.22

- Normal: ES 2.5% is almost equal to its VaR 1% (ratio 1.005); the data: the historical ES 2.5% is 1.5–10.5% above the historical VaR 1% (from the unrounded values)
- For the S&P 500, DAX, BET and Bitcoin, Student-t gives the largest ES among the sensible methods (Bitcoin: 13.37% against 10.84%): with $\hat{\nu} < 3$ its tail is heavier than the data beyond the threshold
- Interpretation: EVT and HS agree at 2.5%; the method matters more for ES than for VaR, because ES depends on the whole tail

SFM_ch10_var_methods

Conditioning on Current Volatility

- Volatility clustering (Chapters 8 and 9): a calm day is followed by calm days, a storm by storms
  - an unconditional VaR is too high in calm years and too low at the start of a crisis
- Conditional model: $r_{t+\mathbf{1}} = \mu + \sigma_{t+\mathbf{1}} z_{t+\mathbf{1}}$, z i.i.d. with mean 0 and variance 1
  - $\mathrm{VaR}_{t+\mathbf{1}} = -(\mu + \sigma_{t+\mathbf{1}}\, q_\alpha(z))$, $\quad \mathrm{ES}_{t+\mathbf{1}} = -\mu + \sigma_{t+\mathbf{1}}\, \mathrm{ES}_\alpha(z)$
  - $q_\alpha(z)$, $\mathrm{ES}_\alpha(z)$: the quantile and the ES of the standardised innovation z
  - two ingredients: the volatility forecast $\sigma_{t+\mathbf{1}}$ and the distribution of z
- Two choices for the distribution of z
  - **GARCH-t**: standardised Student-t with ν estimated with the GARCH(1,1) Bollerslev (1986)
  - **FHS** (filtered historical simulation): the empirical distribution of the standardised residuals $\hat{z}_t = (r_t - \hat{\mu})/\hat{\sigma}_t$ Hull & White (1998); Barone-Adesi, Giannopoulos & Vosper (1999)

Worked Example: the S&P 500 on 18 September 2026

- GARCH(1,1)-t on all data: $\hat{\mu} = 0.079$, $\hat{\alpha} = 0.123$, $\hat{\beta} = 0.872$, $\hat{\nu} = 6.26$; forecast for the next day $\sigma_{t+1} = 0.714\%$
- **GARCH-t**: $q_{0.01}(z) = t_\nu^{-1}(0.01)\sqrt{(\nu-2)/\nu} = -2.557$; $\mathrm{ES}_{2.5\%}(z) = 2.644$
  - VaR 1% $= -(0.079 + 0.714 \times (-2.557)) = 1.75\%$; ES 2.5% $= -0.079 + 0.714 \times 2.644 = 1.81\%$
- **FHS**: the 1% quantile of the 6 714 standardised residuals is -2.787; their ES 2.5% is 2.952
  - VaR 1% $= -(0.079 + 0.714 \times (-2.787)) = 1.91\%$; ES 2.5% $= 2.03\%$
- Interpretation: a calm market: the conditional VaR 1% is about half of the unconditional 3.46%; the standardised residuals have a heavier left tail than the fitted t

SFM_ch10_conditional_var

Next-Day VaR and ES on Six Series

	VaR 1%			ES 2.5%			
	HS (all)	GARCH-t	FHS	HS (all)	GARCH-t	FHS	σ_{t+1}
S&P 500	3.46	1.75	1.91	3.76	1.81	2.03	0.714
DAX	4.20	2.31	2.48	4.26	2.38	2.55	0.945
BET	4.14	4.51	4.81	4.57	4.72	5.04	1.763
Bitcoin	10.54	7.37	7.99	10.84	8.11	8.57	2.829
TLV	5.11	7.18	7.33	5.37	7.53	8.00	2.791
SNP	4.29	5.06	5.39	4.70	5.31	5.65	1.966

- Forecasts for the day after 18 September 2026, in %; HS (all): the unconditional historical values of the previous section
- S&P 500, DAX, Bitcoin: calm, the conditional values are below the unconditional ones
- Interpretation: BET, TLV and SNP: the current volatility is high, so the conditional VaR is above the long-run historical VaR; the same method gives opposite messages on different markets

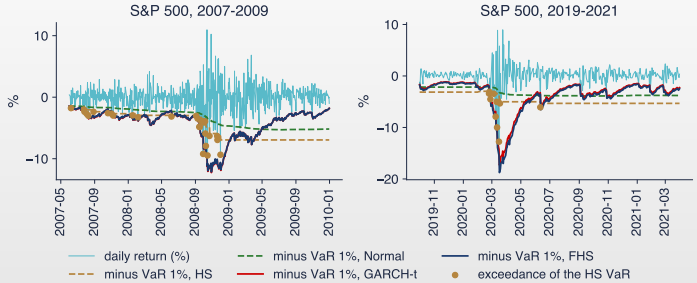
SFM_ch10_conditional_var

Rolling Forecasts: the Design

- **Out of sample**: the forecast for day t uses only data up to day $t-1$
  - S&P 500, DAX, BET: every day since January 2005 (2008 included); Bitcoin: since January 2017
- Four methods
  - **HS** and **Normal**: the last 500 days (about two years)
  - **GARCH-t** and **FHS**: GARCH(1,1)-t re-estimated every 250 days on all past data; σ_t updated every day
- Each day: VaR 1% (for the backtests), VaR 2.5% and ES 2.5% (for the ES backtest)

SFM_ch10_conditional_var

VaR 1% through 2008 and 2020



- S&P 500: daily returns and minus the VaR 1% forecasts of the four methods; dots: exceptions of the HS VaR
- HS and Normal react months late: the 500-day window changes slowly; GARCH-t and FHS jump within days
- Interpretation: after the crisis HS stays high for two years (the ghost effect), while GARCH-t returns to normal levels

SFM_ch10_conditional_var

Stress Periods: Exceptions in 2008 and 2020

	Sep 2008 – Mar 2009					Feb – May 2020				
	HS	Normal	GARCH-t	FHS	exp.	HS	Normal	GARCH-t	FHS	exp.
S&P 500	15	21	5	5	1.5	9	12	3	3	0.7
DAX	6	18	3	3	1.5	8	13	3	3	0.7
BET	8	14	4	4	1.4	5	10	4	4	0.7

- Exceptions of the VaR 1%: days with $r_t < -\mathrm{VaR}_t$; exp.: 1% of the days in the period
- In 2008 the S&P 500 Normal VaR was exceeded on 21 of 146 days, HS on 15; GARCH-t and FHS on 5
- Interpretation: no method is perfect in a crisis, but the conditional ones fail far less: the volatility forecast matters more than the tail shape

SFM_ch10_backtesting

Choosing a VaR Method

Situation	Method	Reason
daily VaR of a liquid position	GARCH-t, FHS	follow the volatility; FHS keeps the empirical tail
long history, quick check	HS	no model, easy to explain
very small α (0.1%), stress tests	EVT	extrapolates beyond the worst day
portfolio of many assets	variance–covariance, copula	dependence; the t copula for joint crashes
moderate skewness and kurtosis	Cornish–Fisher	a quick correction of the Normal

- Never only the Normal for daily data; always report the method, the window and the level
- Whatever the method: backtest it (Section 7)

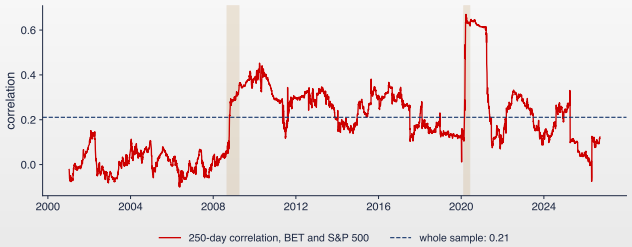
A BET and S&P 500 Portfolio

- Weights $w = (0.5, 0.5)$; simple returns R_t in %, so that $R_{p,t} = w_1 R_{1,t} + w_2 R_{2,t}$ exactly
  - prices joined on the 6 466 common trading days since 2000, then returns
  - each index in its own currency: the currency risk of a RON investor is left out
- Daily data: BET $\hat{\mu} = 0.076$, $\hat{\sigma} = 1.42$; S&P 500 $\hat{\mu} = 0.034$, $\hat{\sigma} = 1.23$; correlation $\hat{\rho} = 0.21$
  - the BVB trades until 17:45 Romanian time and New York opens at 16:30: most of a New York session reaches Bucharest only the next day, so same-day correlation understates the link
- Question: is the portfolio VaR smaller than the average of the two VaRs, and by how much?

The Variance–Covariance Method

- Portfolio: $\mu_p = w^\top \mu$, $\sigma_p^2 = w^\top \Sigma w = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho\sigma_1\sigma_2$ (Σ: covariance matrix)
  - w: the vector of weights; μ: the vector of mean returns; ρ: the correlation of the two assets
  - if returns are jointly Normal: $\mathrm{VaR}_p = -(\mu_p + z_\alpha\sigma_p)$ (the RiskMetrics method J.P. Morgan/Reuters (1996))
- **Step by step**: $\sigma_p^2 = 0.712^2 + 0.614^2 + 2 \times 0.5 \times 0.5 \times 0.369 = 1.068$, $\sigma_p = 1.033$
  - VaR 1% $= -(0.055 - 2.326 \times 1.033) = 2.35\%$; the weighted sum of the two Normal VaRs: $0.5 \times 3.23 + 0.5 \times 2.82 = 3.03\%$
  - **diversification benefit**: 22% of the risk disappears; with $\rho = 1$ it would be zero
- Euler contributions $w_i(\Sigma w)_i/\sigma_p$ ($(\Sigma w)_i$: the i-th entry of Σw) add up to σ_p: BET 56%, S&P 500 44% of the portfolio risk

Correlation in Crisis Periods



- 250-day correlation of BET and S&P 500 daily returns; shaded: the 2008 and 2020 stress periods
- Correlation within each period: 2017 (calm) 0.20; Sep. 2008 – Mar. 2009 0.37; Feb. – May 2020 0.71
- Maximum of the 250-day correlation: 0.67 on 25 March 2020
- Interpretation: diversification is weakest exactly when it is needed: a VaR built on the average correlation is too optimistic in a crisis

SFM_ch10_portfolio_copula

Correlation and Tail Dependence

- Pearson correlation measures linear co-movement over all days; it is dominated by ordinary days Embrechts, McNeil & Straumann (2002)
  - rank correlations (Spearman, Kendall's τ) use only the order of the returns: Kendall $\hat{\tau} = 0.100$, Spearman 0.15
- **Lower tail dependence**: $\lambda_L = \lim_{q \to 0} P(U_1 \leq q \mid U_2 \leq q)$, $U_i = F_i(X_i)$ (the returns on a uniform scale)
  - q: a small probability; F_i: the distribution function of asset i; $\lambda_L \in [0, 1]$
  - the chance that BET has one of its worst days given that the S&P 500 has one of its worst days
  - data: $q = 5\%$: 0.22; $q = 1\%$: 0.19; under independence both would equal q
- Joint crash (both below their own 1% quantile): 0.19% of the days, against 0.01% under independence

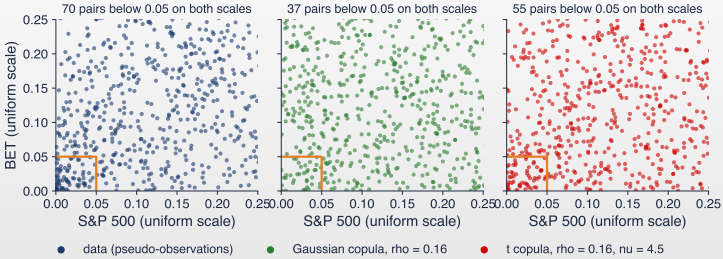
Copulas: an Introduction

- **Sklar's theorem**: every joint distribution function can be written as $F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$ McNeil, Frey & Embrechts (2015)
  - F_1, F_2: the marginal distributions (each asset alone); C: the **copula**, a joint distribution of uniforms on $[0,1]^2$
  - the copula holds all the dependence, the margins all the individual tails
- Practical recipe
  - margins: empirical or fitted (Student-t, EVT); dependence: a copula fitted to the **pseudo-observations** $\hat{U}_{i,t} = \mathrm{rank}(X_{i,t})/(n+1)$
  - simulate (U_1, U_2) from the copula, map back with F_i^{-1}, compute the portfolio return, read VaR and ES
- Textbook: Franke, Härdle & Hafner (2019), Ch. 17

Gaussian and t Copulas

- **Gaussian copula**: the dependence of a bivariate Normal with correlation ρ; $\lambda_L = 0$ for $\rho < 1$
  - extreme days become independent in the limit: joint crashes are rare
- **t copula**: the dependence of a bivariate Student-t with ρ and ν Demarta & McNeil (2005)
  - $\lambda_L = 2\, t_{\nu+1}\Big(-\sqrt{(\nu+1)(1-\rho)/(1+\rho)}\Big) > 0$: crashes cluster across assets
  - $t_{\nu+1}(\cdot)$: the distribution function of a Student-t with $\nu+1$ degrees of freedom; a smaller ν or a larger ρ gives a larger λ_L
- Fit on BET and S&P 500: $\rho = \sin(\pi\hat{\tau}/2) = 0.157$ (valid for both copulas); $\hat{\nu} = 4.50$ by maximum likelihood
  - log-likelihood: Gaussian 92.1, t 219.5; LR $= 254.9$, far above $\chi^2_{0.95}(1) = 3.84$: the t copula wins
  - implied tail dependence $\lambda_L = 0.096$, against 0 for the Gaussian copula

Joint Large Losses: Data and Two Copulas



- Lower-left corner of the uniform scale: pseudo-observations and samples of the same size (6 466) from the fitted copulas
- Days on which both indices are among their worst 5%: data 70, Gaussian 37, t 55; independence: about 16
- Interpretation: the Gaussian copula halves the number of joint bad days; the t copula is much closer to the data

SFM_ch10_portfolio_copula

Copula VaR: the Simulation Step by Step

- **1.** Draw (Z_1, Z_2) from a bivariate Normal with correlation ρ
  - Gaussian copula: $U_i = \Phi(Z_i)$
  - t copula: draw $W \sim \chi^2_\nu$, set $U_i = t_\nu(Z_i\sqrt{\nu/W})$: the common factor $\sqrt{\nu/W}$ makes extremes happen together
- **2.** Map to returns with the empirical quantile functions: $R_i = \hat{F}_i^{-1}(U_i)$
- **3.** Portfolio return $R_p = 0.5R_1 + 0.5R_2$; repeat 200 000 times
- **4.** VaR 1% and ES 2.5% by historical simulation on the simulated R_p
- Same margins in both copulas: any difference in VaR comes from the dependence alone

SFM_ch10_portfolio_copula

Portfolio VaR 1% and ES 2.5% by Method

Method	VaR 1%	ES 2.5%	P(joint crash at 1%)
variance–covariance (Normal)	2.35	2.36	–
Gaussian copula, empirical margins	2.75	2.93	0.03%
t copula, empirical margins	2.81	3.03	0.13%
historical simulation (the data)	3.08	3.30	0.19%

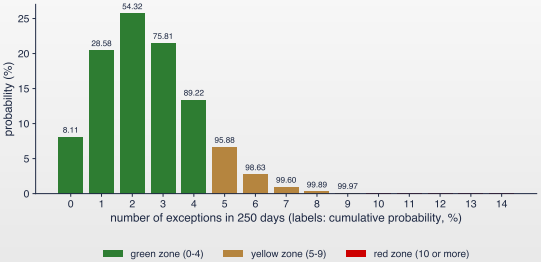
- In %, 50/50 portfolio, 6 466 common days; copulas: 200 000 simulated days
- The Normal method misses both the heavy margins and the tail dependence; the copulas fix the margins, the t copula also the joint crashes
- Interpretation: the t copula still understates the historical VaR: the BET–S&P 500 link is stronger in crises than any constant copula can describe

SFM_ch10_portfolio_copula

The Hit Sequence

- **Exception** (hit): a day with $r_t < -\mathrm{VaR}_t$; $I_t = 1\{r_t < -\mathrm{VaR}_t\}$
  - the VaR must be the forecast made the day before: no look-ahead
- A correct VaR at level α: the I_t are i.i.d. Bernoulli(α) Christoffersen (1998)
  - **unconditional coverage**: $P(I_t = 1) = \alpha$: the right number of exceptions
  - **independence**: an exception today does not make one tomorrow more likely
- Number of exceptions in n days: $x = \sum_t I_t \sim \mathrm{Binomial}(n, \alpha)$, mean $n\alpha$, variance $n\alpha(1-\alpha)$

Exceptions in 250 Days under a Correct VaR 1%



- $\mathrm{Binomial}(250, 0.01)$: expected 2.5 exceptions; 0 exceptions with probability 8.11%, 5 or more with probability $100\% - 89.22\%$, about 11%
- Interpretation: one year of data separates a good model from a bad one only roughly: a correct model lands in the yellow zone about one year in nine

SFM_ch10_backtesting

Kupiec's POF Test

- **POF** (proportion of failures) Kupiec (1995): H_0: $P(I_t = 1) = \alpha$; $\hat{\pi} = x/n$
  - $LR_{uc} = -2\left[(n-x)\ln(1-\alpha) + x\ln\alpha - (n-x)\ln(1-\hat{\pi}) - x\ln\hat{\pi}\right] \sim \chi^2(1)$ under H_0
  - x: exceptions in n days; $\hat{\pi}$: the observed rate; LR_{uc} compares the likelihood under the rate α with that under $\hat{\pi}$
  - $LR_{uc} \geq 0$; a likelihood-ratio test (Chapter 6); reject at 5% if $LR_{uc} > 3.84$
- **Step by step**: $n = 250$, $x = 7$, $\hat{\pi} = 0.028$
  - $\ln L_0 = 243\ln 0.99 + 7\ln 0.01 = -34.68$; $\ln L_1 = 243\ln 0.972 + 7\ln 0.028 = -31.93$
  - $LR_{uc} = -2(-34.68 - (-31.93)) = 5.50$, p $= 0.019$: rejected at 5%
- Low power with one year of data
  - in 250 days, any count from 1 to 6 is accepted at 5%: a model with a true rate of 2% passes with probability 76%
  - too few exceptions are also rejected: an overly prudent VaR wastes capital

Kupiec on 21 Years of S&P 500 Forecasts

- GARCH-t VaR 1%, 3 January 2005 – 18 September 2026: $n = 5\,460$ days, $x = 95$ exceptions, $\hat{\pi} = 0.0174$
  - expected $n\alpha = 55$; a 95% binomial range for the rate under H_0: 0.75–1.26%
- **Step by step**
  - $\ln L_0 = 5365 \ln 0.99 + 95 \ln 0.01 = -491.4$
  - $\ln L_1 = 5365 \ln(1 - 0.0174) + 95 \ln 0.0174 = -479.0$
  - $LR_{uc} = -2(\ln L_0 - \ln L_1) = 24.73$, $p < 0.001$
- Interpretation: with 21 years of data even a rate of 1.74% instead of 1% is clearly rejected; the missing leverage effect (Chapter 9) is a likely cause

SFM_ch10_backtesting

Christoffersen's Independence Test

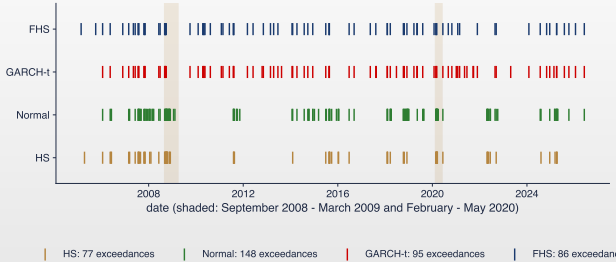
- Count the transitions: n_{ij} = number of days with $I_{t-1} = i$ and $I_t = j$; $\hat{\pi}_{01} = \dfrac{n_{01}}{n_{00} + n_{01}}$, $\hat{\pi}_{11} = \dfrac{n_{11}}{n_{10} + n_{11}}$
  - $\hat{\pi}_{01}$: the rate of exceptions after a day without one; $\hat{\pi}_{11}$: after a day with one
  - H_0: $\pi_{01} = \pi_{11}$ (an exception today does not change the chance of one tomorrow) Christoffersen (1998)
  - $LR_{ind} = -2 \ln \dfrac{(1-\hat{\pi})^{n_{00}+n_{10}} \hat{\pi}^{n_{01}+n_{11}}}{(1-\hat{\pi}_{01})^{n_{00}} \hat{\pi}_{01}^{n_{01}} (1-\hat{\pi}_{11})^{n_{10}} \hat{\pi}_{11}^{n_{11}}} \sim \chi^2(1)$
- **Conditional coverage**: $LR_{cc} = LR_{uc} + LR_{ind} \sim \chi^2(2)$; reject at 5% if $LR_{cc} > 5.99$
- **Example**: S&P 500, HS VaR 1%, 5 460 days: $n_{00} = 5313$, $n_{01} = 69$, $n_{10} = 69$, $n_{11} = 8$
  - $\hat{\pi}_{01} = 1.28\%$, $\hat{\pi}_{11} = 10.4\%$: after an exception the next one is 8 times more likely
  - $LR_{ind} = 19.43$ ($p < 0.001$): the exceptions cluster

The Basel Traffic Light

x	Zone	Cum. (%)	Plus	Factor
0	green	8.11	0.00	3.00
1	green	28.58	0.00	3.00
2	green	54.32	0.00	3.00
3	green	75.81	0.00	3.00
4	green	89.22	0.00	3.00
5	yellow	95.88	0.40	3.40
6	yellow	98.63	0.50	3.50
7	yellow	99.60	0.65	3.65
8	yellow	99.89	0.75	3.75
9	yellow	99.97	0.85	3.85
≥ 10	red	99.99	1.00	4.00

- BCBS (1996b): count the exceptions of the one-day VaR 1% over the last 250 days
  - green: no action; yellow: the multiplier grows by the plus factor; red: the model is presumed wrong
- Capital $= \max\{\mathrm{VaR}_{t-\mathbf{1}},\ (3+\text{plus}) \times \overline{\mathrm{VaR}}_{\mathbf{60}}\}$: the larger of yesterday's 10-day VaR 1% and the multiplier times its 60-day average
  - zones from the binomial: yellow starts where the cumulative probability first exceeds 95%, red where it reaches 99.99%
- x: exceptions in 250 days; Cumul.: $P(X \leq x)$ under $\mathrm{Binomial}(250, 0.01)$
- FRTB keeps the count of VaR 1% exceptions at bank and desk level (MAR32, BCBS (2023a))

Exceptions over Time



- S&P 500, 2005–2026: each mark is a day on which the VaR 1% of the method was exceeded; expected about 55 per method
- Interpretation: HS and Normal exceptions come in bursts (2007–2008, 2011, 2022); GARCH-t and FHS exceptions are spread over time, but there are more of them in calm years

SFM_ch10_backtesting

Backtest Results: S&P 500 and DAX

	Method	x	rate (%)	p POF	p ind.	p c.c.	max 250 d.	Z_2 (p)
S&P 500	HS	77	1.41	0.004	< 0.001	< 0.001	18	-0.39 (< 0.001)
	Normal	148	2.71	< 0.001	< 0.001	< 0.001	29	-1.15 (< 0.001)
	GARCH-t	95	1.74	< 0.001	0.114	< 0.001	10	-0.54 (< 0.001)
	FHS	86	1.58	< 0.001	0.060	< 0.001	10	-0.20 (0.013)
DAX	HS	67	1.21	0.121	< 0.001	< 0.001	12	-0.26 (0.002)
	Normal	133	2.41	< 0.001	< 0.001	< 0.001	19	-0.87 (< 0.001)
	GARCH-t	84	1.52	< 0.001	0.549	0.001	9	-0.49 (< 0.001)
	FHS	77	1.40	0.005	0.418	0.015	9	-0.18 (0.014)

- VaR 1% since 3 January 2005 (about 5 500 days); x: exceptions; p ind.: Christoffersen independence; p c.c.: conditional coverage; max 250 d.: the worst 250-day count; Z_2: the ES backtest at the end of this section
- Normal: rejected everywhere; HS: exceptions cluster; GARCH-t and FHS: independent exceptions, but a rate of 1.4–1.7%

SFM_ch10_backtesting

Backtest Results: BET and Bitcoin

	Method	x	rate (%)	p POF	p ind.	p c.c.	max 250 d.	Z_2 (p)
BET	HS	66	1.22	0.123	0.001	0.002	14	−0.27 (0.001)
	Normal	122	2.25	< 0.001	< 0.001	< 0.001	21	−0.83 (< 0.001)
	GARCH-t	65	1.20	0.158	0.807	0.358	9	−0.22 (0.007)
	FHS	65	1.20	0.158	0.240	0.185	7	−0.18 (0.021)
Bitcoin	HS	33	0.93	0.672	0.316	0.553	6	−0.07 (0.224)
	Normal	69	1.94	< 0.001	0.057	< 0.001	16	−0.56 (< 0.001)
	GARCH-t	48	1.35	0.045	0.683	0.123	8	−0.41 (< 0.001)
	FHS	24	0.68	0.040	0.155	0.044	6	0.12 (0.834)

- VaR 1% since 3 January 2005 (Bitcoin since 1 January 2017); the columns as on the previous slide
- BET: GARCH-t and FHS pass all three tests; Bitcoin: HS passes, FHS is too prudent (0.68%)

SFM_ch10_backtesting

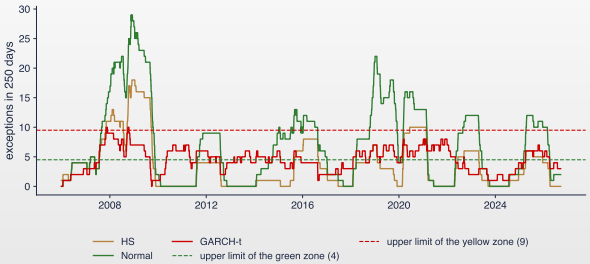
Ranking of the Forecasts and the Current Traffic-Light Zone

	average quantile loss $\times 100$				exceptions, last 250 days			
	HS	Normal	GARCH-t	FHS	HS	Normal	GARCH-t	FHS
S&P 500	5.17	5.80	3.71	3.68	0	2	3	3
DAX	5.15	5.52	4.08	4.05	1	3	3	3
BET	6.04	6.26	4.59	4.62	4	6	3	2
Bitcoin	13.57	14.52	13.53	13.75	3	4	2	2

- Tests say whether a model is acceptable; a **scoring function** says which of two models is better
  - VaR is elicitable: the average quantile loss $\frac{1}{n}\sum_t(\alpha - I_t)(r_t + \mathrm{VaR}_t)$ ranks the forecasts Gneiting (2011)
- The equity indices: GARCH-t and FHS have the smallest loss, the Normal the largest; Bitcoin: the four methods are close
- Interpretation: on the last 250 days every model is in the green zone except the Normal VaR of the BET (6, yellow): a recent window says little about a model

SFM_ch10_backtesting

The Traffic Light through Time



- S&P 500: number of VaR 1% exceptions in the last 250 days, for every day from 28 December 2005 (the end of the first 250-day window); dashed: the zone limits
- Share of days in the red zone: Normal 33.3%, HS 11.8%, GARCH-t 0.5%; worst count: 29, 18 and 10
- Interpretation: a bank using the Normal or HS VaR would have paid the red-zone penalty for most of 2008–2009; GARCH-t reaches it on only 0.5% of the days

SFM_ch10_backtesting

Backtesting ES: Acerbi and Székely

- ES is not elicitable alone, but it can be tested together with VaR Acerbi & Székely (2014)
  - $Z_2 = \frac{1}{T\alpha} \sum_{t=1}^{T} \frac{r_t I_t}{\mathrm{ES}_t} + 1, \quad I_t = 1\{r_t < -\mathrm{VaR}_t\}$, $\alpha = 2.5\%$
  - T: the number of test days; VaR_t, ES_t: the forecasts for day t; each exception adds r_t/ES_t
  - if VaR and ES are right, $E[Z_2] = 0$; $Z_2 < 0$: the losses beyond VaR are larger than the forecast ES, or too frequent
- p-value by simulation: draw returns from each day's forecast distribution, recompute Z_2 1000 times
  - the values of Z_2 and their p-values: the last column of the two backtest tables
- Interpretation: all four ES 2.5% forecasts are too small for the equity indices; FHS is closest, because it takes the tail of z from the data

SFM_ch10_backtesting

Case Study: VaR Models at Commercial Banks

- Berkowitz & O'Brien (2002): the first study of the VaR forecasts that large US banks actually reported, with their daily trading P&L
  - design: the banks' VaR 1% against the realised P&L; a simple GARCH model of each bank's P&L as the benchmark
- Findings
  - the banks' VaR was conservative on average, yet the exceptions that did occur came in clusters
  - the GARCH benchmark gave lower VaRs with comparable coverage: it followed changes in volatility better
- The same pattern as in our data: window-based VaR is either too high in calm times or too slow in storms; conditional VaR adapts
- Further reading on forecasting the quantile directly: CAViaR Engle & Manganelli (2004)

An Open Research Question

- **Does the tail dependence between the BVB and world markets jump in crises, and does it come back?**
  - today: correlation 0.20 in a calm year, 0.71 in spring 2020; a constant t copula gives $\lambda_L = 0.096$
  - a portfolio VaR with a constant copula is still below the historical one (Section 6)
- Why it is open: crises are few, the BVB trades at other hours, and tail estimates need many observations
- Related work on entropy as a risk measure, as further reading: Pele, Lazar & Dufour (2017); Pele & Mazurencu-Marinescu-Pele (2019)
- AI tools can speed up such a study, but its results must still be checked Wang et al. (2023)

Possible Uses of AI Tools

- **Literature**: a list of studies on time-varying copulas and on contagion between emerging and developed markets
- **Code**: a first draft of rolling t-copula estimates (250 days) for BET and S&P 500, with ρ, ν and λ_L over time
- **Robustness**: two-day returns (different trading hours), DAX instead of S&P 500, other windows
- Example prompt
  - `Write Python code that fits a bivariate t copula by maximum likelihood to the pseudo-observations of daily BET and S&P 500 returns on rolling windows of 250 days moved by 21 days, and plots rho, nu and the lower tail dependence over time.`

What to Check

- The VaR convention: "VaR 1%" is the loss exceeded with probability 1%, a positive number; an answer that names the complementary probability or reports a negative VaR mixes conventions
- Prices joined on common days before computing returns; the BVB and New York trade at different hours
- Pseudo-observations: ranks divided by $n + 1$, not by n (otherwise the copula density is infinite at 1)
- Tail dependence from 250 days rests on a handful of joint bad days: report its uncertainty (bootstrap, Chapter 2)
- Overlapping windows: consecutive estimates are not independent evidence
- References: every cited paper must exist; check the DOI

Project Idea

- **Question**: does a time-varying copula give a better portfolio VaR for a BVB and world-equity portfolio than a constant one?
  - data: BET, BET-TR, S&P 500, DAX since 2005 (EODHD)
- Steps
  - GARCH-t margins for each index (Chapter 9); constant and rolling t copula for the dependence
  - one-day VaR 1% and ES 2.5% of a 50/50 portfolio, out of sample since 2008
  - Kupiec, Christoffersen and Z_2 for both versions, with special attention to 2008, 2020 and 2022
- Deliverable: one table, one chart, and a paragraph on what the data can and cannot show
- Declare any AI use, and list the errors of the AI that you corrected

Key Takeaways

- VaR 1% $= -q_{1\%}$: the loss exceeded on one day in a hundred; ES 2.5%: the average loss on the worst 2.5% of days
- ES is coherent; VaR can penalise diversification; VaR is elicitable, ES only together with VaR
- The Normal VaR is 14–24% too low on daily data; HS, Student-t and EVT agree at 1%; Cornish–Fisher fails for large kurtosis
- Conditional VaR (GARCH-t, FHS) follows the storms and passes backtests that window methods fail
- Portfolio risk: correlation rises in crises; tail dependence (t copula) matters for joint crashes
- Backtest with Kupiec (rate), Christoffersen (clustering), the traffic light (regulation) and Z_2 (ES)

Key Formulas

Quantity	Formula
VaR	$\mathrm{VaR}_\alpha = -q_\alpha(X)$; HS: $-r_{(\lceil n\alpha \rceil)}$
ES	$\mathrm{ES}_\alpha = -\frac{1}{\alpha}\int_0^\alpha q_u\,du$; HS: $-\frac{1}{k}\sum_{i\le k} r_{(i)}$
Normal	$-(\mu+\sigma z_\alpha)$, $-\mu+\sigma\varphi(z_\alpha)/\alpha$
EVT (POT)	$u+\frac{\beta}{\xi}[(n\alpha/N_u)^{-\xi}-1]$, $\mathrm{ES}=(\mathrm{VaR}+\beta-\xi u)/(1-\xi)$
Conditional	$-(\mu+\sigma_{t+1}q_\alpha(z))$
Portfolio	$\sigma_p^2 = w^\top \Sigma w$
t copula	$\lambda_L = 2t_{\nu+1}(-\sqrt{(\nu+1)(1-\rho)/(1+\rho)})$
Kupiec	$LR_{uc} = -2\ln[(1-\alpha)^{n-x}\alpha^x/((1-\hat\pi)^{n-x}\hat\pi^x)] \sim \chi^2(1)$
Christoffersen	$LR_{cc} = LR_{uc} + LR_{ind} \sim \chi^2(2)$
Acerbi–Székely	$Z_2 = \frac{1}{T\alpha}\sum_t r_t I_t/\mathrm{ES}_t + 1$

Self-Assessment

- **Question**: a desk reports "VaR 1% $= -2.4\%$". What is wrong?
  - **Answer**: the sign: -2.4% is the 1% quantile $q_{0.01}$; VaR 1% $= -q_{0.01} = 2.4\%$, a loss
- **Question**: a model has 3 exceptions in 250 days, all in the same week. Which test detects the problem?
  - **Answer**: not Kupiec (3 is close to 2.5); Christoffersen's independence test can, if the exceptions fall on consecutive days (it looks only at day-to-day transitions), and with 3 exceptions its power is low
- **Question**: two portfolios have the same VaR 1%. Can their ES 2.5% differ?
  - **Answer**: yes: ES depends on the whole tail beyond the quantile, VaR only on one point
- Next: Chapter 11, the fractal markets hypothesis and long memory

References (1/2)

- Acerbi, C., Székely, B. (2014). Backtesting expected shortfall. MSCI Research; published in *Risk*, December 2014.
- Acerbi, C., Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7), 1487–1503.
- Artzner, P., Delbaen, F., Eber, J.-M., Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3), 203–228.
- Barone-Adesi, G., Giannopoulos, K., Vosper, L. (1999). VaR without correlations for portfolios of derivative securities. *Journal of Futures Markets*, 19(5), 583–602.
- Basel Committee on Banking Supervision (1996a). Amendment to the capital accord to incorporate market risks. Bank for International Settlements.
- Basel Committee on Banking Supervision (1996b). Supervisory framework for the use of "backtesting" in conjunction with the internal models approach to market risk capital requirements. Bank for International Settlements.
- Basel Committee on Banking Supervision (2019). Minimum capital requirements for market risk. Bank for International Settlements.
- Basel Committee on Banking Supervision (2023a). MAR32: Internal models approach: backtesting and P&L attribution test requirements. Basel Framework, version effective 1 January 2023. Bank for International Settlements.
- Basel Committee on Banking Supervision (2023b). MAR33: Internal models approach: capital requirements calculation. Basel Framework, version effective 1 January 2023. Bank for International Settlements.
- Berkowitz, J., O'Brien, J. (2002). How accurate are value-at-risk models at commercial banks? *Journal of Finance*, 57(3), 1093–1111.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.
- Borak, S., Härdle, W.K., López-Cabrera, B. (2013). *Statistics of Financial Markets: Exercises and Solutions*, 2nd ed. Springer.
- Christoffersen, P.F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4), 841–862.
- Cornish, E.A., Fisher, R.A. (1938). Moments and cumulants in the specification of distributions. *Revue de l'Institut International de Statistique*, 5(4), 307–320.
- Daníelsson, J., Jorgensen, B.N., Samorodnitsky, G., Sarma, M., de Vries, C.G. (2013). Fat tails, VaR and subadditivity. *Journal of Econometrics*, 172(2), 283–291.

References (2/2)

- Demarta, S., McNeil, A.J. (2005). The t copula and related copulas. *International Statistical Review*, 73(1), 111–129.
- Embrechts, P., McNeil, A.J., Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In M. Dempster (ed.), *Risk Management: Value at Risk and Beyond*, 176–223. Cambridge University Press.
- Engle, R.F., Manganelli, S. (2004). CAViaR: conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4), 367–381.
- Fissler, T., Ziegel, J.F. (2016). Higher order elicitability and Osband's principle. *Annals of Statistics*, 44(4), 1680–1707.
- Franke, J., Härdle, W.K., Hafner, C.M. (2019). *Statistics of Financial Markets: An Introduction*, 5th ed. Springer.
- Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106(494), 746–762.
- Hull, J., White, A. (1998). Incorporating volatility updating into the historical simulation method for value-at-risk. *Journal of Risk*, 1(1), 5–19.
- J.P. Morgan/Reuters (1996). *RiskMetrics – Technical Document*, 4th ed. New York.
- Kupiec, P.H. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 3(2), 73–84.
- McNeil, A.J., Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7(3–4), 271–300.
- McNeil, A.J., Frey, R., Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools*, revised ed. Princeton University Press.
- Pele, D.T., Lazar, E., Dufour, A. (2017). Information entropy and measures of market risk. *Entropy*, 19(5), 226.
- Pele, D.T., Mazurencu-Marinescu-Pele, M. (2019). Using high-frequency entropy to forecast Bitcoin's daily value at risk. *Entropy*, 21(2), 102.
- Wang, H., Fu, T., Du, Y., Gao, W., et al. (2023). Scientific discovery in the age of artificial intelligence. *Nature*, 620, 47–60.