Project 1: Market Risk – VaR, ES, and Copulas

Deadline: November 28, 2025

Contact: Andrea Ruglioni (`andrea.ruglioni@epfl.ch`)

General instructions. You may work in groups of up to three students (smaller groups are also allowed). Your submission must contain:

1. A single **PDF report** including a clear description of your methodology, the results of your experiments (tables and figures), and your *commented code* in an appendix.
2. The corresponding **code files** in Python (.py or .ipynb).

The PDF should be fully self-contained: all results, plots, and tables must be visible in the report without executing the code. When the provided scripts are executed, they should exactly reproduce the results shown in your PDF. Please list all group members on the title page of your report.

Grading. The project is worth 15% of the final grade. You will be evaluated based on the correctness of your implementation, the clarity and completeness of your explanations, and the quality of your analysis and interpretation of results.

Project description. This project focuses on market risk measurement using Value-at-Risk (VaR) and Expected Shortfall (ES), and on modeling cross-asset dependence with copulas. You will work with *historical financial data*. You can use the code in the provided template, `Project1_template.ipynb`, to fetch and prepare data.

Useful Python packages include `numpy`, `pandas`, `scipy` and `statsmodels` (for correlations and distribution fitting), `arch`, `matplotlib`, `seaborn`, and `yfinance`. For copulas, you may use the `copulae` package or implement fitting/simulation yourself. You are free to use other packages as needed.

1. **Empirical stylized facts.** Use the following three stocks: **AAPL, META, JPM**, over the recent window January 1, 2023–June 30, 2025. Fetch daily closing prices (adjusted for splits/dividends) using `yfinance` in Python through the provided template.
   a) Construct *log-returns* $R_t = \log(P_t/P_{t-1}), t = 1, \dots, T$; handle missing data appropriately if needed. Plot return series; comment on trends/volatility.
   b) Compute and plot cross-correlation and autocorrelation functions for both raw returns and absolute returns from a), that is $\mathrm{Corr}(R_{t,n}, R_{t-h,m})$ and $\mathrm{Corr}(|R_{t,n}|, |R_{t-h,m}|)$ for $n, m \in \{\text{AAPL, META, JPM}\}$ and lags $h = 0, 1, \dots, 25$. Discuss the evidence of volatility clustering, linear dependence, and cross-asset linkages.
   c) Graphically assess the normality of returns using QQ plots against the normal distribution. Then, introduce and apply the Jarque-Bera test to support your findings.

2. **Single-window modeling: VaR, ES, and distributions.** Analyze each asset separately on the first estimation window of length $W = 252$, i.e., approximately one trading year. Clearly describe your fitting procedure and any choices made. Use the loss convention $L_t = -R_t$, for $t = 1, \dots, W$ (i.e., right tail is risky). *Fit and compare* the following models and report the estimated 1-day ahead VaR and ES at levels $\alpha = 95\%$ and $\alpha = 99\%$ To compare different fitted loss distributions, plot their estimated probability density functions (pdfs) in a single plot. This allows you to visually assess the fit and tail behavior.

a) *Historical simulation:* obtain the empirical cdf and use it to compute the VaR/ES of L.

b) *Gaussian:* fit (μ, σ) and use the closed-form solutions for VaR/ES.

c) *Student-t:* fit (ν, μ, σ) using maximum likelihood estimation (MLE) and use the closed-form solutions for VaR/ES; discuss the impact of degrees of freedom on tails.

d) *Conditional parametric:* the previous models assume i.i.d. losses, now we make use of conditional models. The mean is defined through an autoregressive model AR(p) where you choose p based on the autocorrelation function (ACF) from 1.b) and the partial ACF. The volatility is modeled through a GARCH(1,1) model, with Gaussian innovations. Use the fitted model to obtain 1-step-ahead VaR/ES forecasts based on the conditional mean $\hat{\mu}_{W+1}$ and volatility $\hat{\sigma}_{W+1}$ estimates from the AR(p) and GARCH(1,1) parts, respectively.

e) *Filtered Historical Simulation (FHS):* describe and implement the FHS procedure following Barone-Adesi et al. (1999) or a similar reference. Using the fitted AR(p) + GARCH(1,1) approach from 2.d), obtain the model's residuals $\hat{\epsilon}_t$ and standardized residuals $\tilde{\epsilon}_t = \hat{\epsilon}_t/\hat{\sigma}_t$, $t = 1, \ldots, W$. Instead of assuming a parametric distribution for the innovation $\tilde{\epsilon}_{W+1}$ (such as in 2.d), employ the non-parametric bootstrap (i.e., resampling $M = 1000$ values with replacement) from the constructed sample of $\tilde{\epsilon}$ to obtain the 1-step-ahead VaR/ES forecasts.

3. **Backtesting VaR and ES.** Use a rolling window of size $W = 252$ to produce 1-step-ahead forecasts of VaR/ES at the confidence levels $\alpha = 95\%$ and $\alpha = 99\%$ for each asset and each method from 2.a)–2.e) over the following *out-of-sample period.* Clearly describe your backtesting (i.e., rolling-window) procedure: at each time t, fit the models on $\{t - W + 1, \ldots, t\}$, forecast at $t + 1$, advance by one day, and repeat until $t + 1 = T$, where T is the end of your sample. Plot the VaR forecasts along with the realized returns. Unless otherwise stated, use 95% as the primary confidence level.

   a) Implement and apply the Kupiec (1995) *Proportion-Of-Failures* (POF) test (unconditional coverage) and the Christoffersen (1998) conditional coverage test (joint test of correct exception rate and independence). Report test statistics, p-values, and discuss which models pass/fail. Interpret whether failures are due to biased coverage or clustered exceptions.

   Refer to the referenced papers for further details, and go over the brief descriptions in the Appendix.

   b) *ES backtest:* Describe and apply the Acerbi and Székely (2014) Z_1 test for ES backtesting with $M = 1000$ simulations. State clearly the null and alternative hypotheses, emphasizing that the test is primarily one-sided (detecting underestimation of ES). Discuss which models pass/fail, report the corresponding p-values, and interpret the results.

   Refer to the referenced paper for further details, and go over the brief descriptions in the Appendix.

4. **Copula fitting.** On the single estimation window of length $W = 252$ (i.e., same as in question 2), model cross-asset *dependence* across AAPL, META, and JPM using copulas.

   a) For each asset i, let $U_{t,i} = \frac{\text{rank}(R_{t,i})}{W+1}$, $t = 1, \ldots, W$ be the empirical quantiles, sometimes known as pseudo-observations. Plot dependence between asset pairs both in the raw returns space (i.e., via the scatter plot of $(R_{t,i}, R_{t,j})$) and the pseudo-observations space (i.e., via the scatter plot of $(U_{t,i}, U_{t,j})$). Comment on tail dependence, symmetry/asymmetry, and how these features motivate the choice of copula.

   b) Fit a Gaussian copula and a Student-t copula and explain the fitting procedure. Report the obtained parameter estimates.

c) Explain how to simulate from a copula, and use it to generate synthetic data. Thereby, sample T points from the fitted copulas, with T being the total number of observations in the dataset. Using the inverse empirical cdf of each asset, transform the simulated uniform samples into simulated returns. Visually compare the original and simulated returns, and describe what you observe.

5. **Backtesting portfolio VaR and ES.** Construct an *equal-weighted* portfolio of AAPL, META, and JPM. Use the same univariate methods as in question 2 and include also the copula-based approach from question 4. Estimate 1-day ahead VaR/ES at $\alpha = 95\%$ and $\alpha = 99\%$, then *backtest and compare* them using rolling-window forecasts as in question 3. Compare portfolio VaR/ES backtests across univariate and copula-based approaches. Assess whether explicit dependence modeling improves accuracy, especially in the tails. Discuss potential reasons.

   a) *Univariate models:* use the same backtesting procedure as in question 3, but now on the portfolio returns.

   b) *Copulas:* at each time t, fit the copulas as in 4.b) on the last W days of portfolio component returns. Simulate $N = 1000$ returns from each fitted copula, then estimate VaR/ES from the simulated portfolio returns.

   c) *Backtesting and comparison:* present a compact comparison of backtests. Does dependence modeling improve results? Why/why not?

References

C. Acerbi and B. Székely. Backtesting expected shortfall. *Risk*, 27(11):76–81, 2014.

Giovanni Barone-Adesi, Kostas Giannopoulos, and Les Vosper. VaR without correlations for portfolios of derivative securities. *The Journal of Futures Markets*, 19(5):583–602, 1999.

P. F. Christoffersen. Evaluating interval forecasts. *International Economic Review*, 39(4):841–862, 1998.

P. H. Kupiec. Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, 3(2):73–84, 1995.

A Backtesting VaR and ES

A.1 Kupiec Proportion-of-Failures (POF) Test

Kupiec (1995) test evaluates whether the observed frequency of VaR violations is consistent with the nominal level. It addresses *unconditional coverage* only, not the timing of violations.

Let

$$I_t = \mathbf{1}\{L_t > \hat{q}_{t,\alpha}\}, \quad t = 1, \dots, T,$$

be the indicator of a VaR violation, where L_t denotes the realized loss and $\hat{q}_{t,\alpha}$ the one-step-ahead VaR forecast at level α. The total number of violations is

$$N = \sum_{t=1}^{T} I_t, \qquad \hat{p} = \frac{N}{T}.$$

Hypotheses: under the null, $\{I_t\}$ are i.i.d. Bernoulli(p) with $p = 1 - \alpha$. The alternative is that the true violation probability differs from p.

Likelihoods:

$$L_0 = (1-p)^{T-N} p^N,$$
$$L_1 = (1-\hat{p})^{T-N} \hat{p}^N.$$

Test statistic:

$$LR_{\text{POF}} = -2\log\left(\frac{L_0}{L_1}\right) = -2\log\left(\frac{(1-p)^{T-N} p^N}{(1-\hat{p})^{T-N}\hat{p}^N}\right).$$

As $T \to \infty$, $LR_{\text{POF}} \sim \chi^2_1$. Large values indicate that the model produces too many or too few exceptions relative to the nominal level.

Implementation steps:

1. Compute I_t for each t.
2. Count $N = \sum_t I_t$, compute $\hat{p} = N/T$.
3. Plug into the likelihood ratio formula.
4. Compare LR_{POF} to the χ^2_1 distribution (or report the p-value).
5. Interpret: reject if the observed exception rate is inconsistent with $p = 1 - \alpha$.

A.2 Christoffersen Independence and Conditional Coverage Tests

The POF test ignores whether violations cluster in time. Christoffersen (1998) extends backtesting by testing for independence of exceptions, and combining both coverage and independence into a *conditional coverage* test.

Let $\{I_t\}$ be the violation sequence as above. Model it as a first-order Markov chain with transition probabilities

$$\pi_{ij} = \Pr(I_t = j \mid I_{t-1} = i), \quad i, j \in \{0, 1\}.$$

Transition counts:

$$N_{ij} = \#\{t : I_{t-1} = i,\, I_t = j\}.$$

Likelihoods:

$$L_1 = (1-\hat{\pi}_0)^{N_{00}} \hat{\pi}_0^{N_{01}} (1-\hat{\pi}_1)^{N_{10}} \hat{\pi}_1^{N_{11}},$$
$$L_0 = (1-\hat{p})^{N_{00}+N_{10}} \hat{p}^{N_{01}+N_{11}},$$

where

$$\hat{\pi}_0 = \frac{N_{01}}{N_{00}+N_{01}}, \quad \hat{\pi}_1 = \frac{N_{11}}{N_{10}+N_{11}}, \quad \hat{p} = \frac{N_{01}+N_{11}}{N_{00}+N_{01}+N_{10}+N_{11}}.$$

Independence test:

$$LR_{\text{ind}} = -2\log\left(\frac{L_0}{L_1}\right) \sim \chi_1^2.$$

Rejection indicates that violations are not independent, but clustered.

Conditional coverage test:

$$LR_{\text{cc}} = LR_{\text{POF}} + LR_{\text{ind}} \sim \chi_2^2.$$

This jointly tests correct coverage and independence. Rejection indicates that the sequence $\{I_t\}$ is not i.i.d. Bernoulli(p).

Implementation steps:

1. Build the violation sequence I_t.
2. Count transitions $N_{00}, N_{01}, N_{10}, N_{11}$.
3. Estimate $\hat{\pi}_0, \hat{\pi}_1, \hat{p}$.
4. Compute likelihoods L_0 and L_1.
5. Calculate LR_{ind}.
6. Compute $LR_{\text{cc}} = LR_{\text{POF}} + LR_{\text{ind}}$.
7. Compare to the χ^2 distributions (df=1 for independence, df=2 for joint).
8. Interpret results:
   - Fail unconditional coverage $\Rightarrow$ too many/few violations.
   - Fail independence $\Rightarrow$ clustering of violations.
   - Fail conditional coverage $\Rightarrow$ model inadequate overall.

A.3 Acerbi–Székely Z_1 Test

Unlike VaR, ES cannot be backtested by simple exception counting. Acerbi and Székely (2014) propose a family of nonparametric ES backtests. The simplest of these is the Z_1 test, which evaluates whether realized tail losses are consistent with the ES forecasts, given that the corresponding VaR forecasts at level α are correct.

Let L_t denote the one-step-ahead loss, and let $(\hat{q}_{t,\alpha}, \hat{e}_{t,\alpha})$ be the model's forecasted VaR and ES pair. Define the exception indicator and the total number of violations:

$$I_t = \mathbf{1}\{L_t > \hat{q}_{t,\alpha}\}, \qquad N = \sum_{t=1}^{T} I_t.$$

Null hypothesis: if both VaR and ES are correctly specified, the following conditional expectation identity holds:

$$\mathbb{E}\left[\frac{L_t}{\hat{e}_{t,\alpha}} - 1 \,\middle|\, I_t = 1\right] = 0.$$

Test statistic: replace the conditional expectation with the empirical average over realized violations:

$$Z_1 = \frac{1}{N}\sum_{t=1}^{T} \frac{L_t I_t}{\hat{e}_{t,\alpha}} - 1.$$

If $N = 0$, there are no tail events and the test is inconclusive. Otherwise, Z_1 is a *single number* computed from the data. Under the null, Z_1 should be close to zero. Positive values ($Z_1 > 0$) indicate that realized tail losses are, on average, larger than the ES forecasts (underestimation of risk).

Why simulation is needed: Unlike the VaR tests, the sampling distribution of Z_1 cannot be expressed in closed form; it depends on the model's predictive distribution. To obtain a reference distribution, Acerbi and Székely propose a simulation approach:

1. For each forecast date t, generate M artificial losses $\{L_t^{(m)}\}_{m=1}^{M}$ from the model's one-step predictive distribution (the same one used to compute $\hat{q}_{t,\alpha}$ and $\hat{e}_{t,\alpha}$).
2. For each simulated path m, compute violation indicators $I_t^{(m)}$, the number of exceptions $N^{(m)}$, and the corresponding statistic
$$Z_1^{(m)} = \frac{1}{N^{(m)}}\sum_{t=1}^{T} \frac{L_t^{(m)} I_t^{(m)}}{\hat{e}_{t,\alpha}} - 1.$$
Paths with $N^{(m)} = 0$ are discarded.
3. This yields M values $\{Z_1^{(m)}\}$, which approximate the *null distribution* of Z_1 under the model.
4. Compare the observed Z_1 (from real data) to this simulated distribution. The one-sided p-value is computed as
$$p = \frac{1}{M}\sum_{m=1}^{M} \mathbf{1}\{Z_1^{(m)} \geq Z_1\}.$$

Interpretation:

- Z_1 (observed) is the realized test statistic from data.
- $\{Z_1^{(m)}\}$ are many simulated test statistics, showing what Z_1 would look like if the ES model were correct.
- If Z_1 is unusually large compared to its simulated null distribution (i.e., p is small), we reject the ES model: realized tail losses are heavier than predicted.

Implementation steps:

1. Compute I_t, N, and the observed statistic Z_1 from data.
2. Simulate M predictive scenarios from the model, compute $\{Z_1^{(m)}\}$.
3. Approximate the null distribution of Z_1 with $\{Z_1^{(m)}\}$.
4. Calculate the one-sided p-value as above.
5. Conclude: reject H_0 if p is below the chosen significance level.