Model validation of the value at risk calculations

– two-asset portfolio case study

Anna Nakonieczna

March, 2025

Contents

Model validation in finance

→ assessing the accuracy, robustness, and reliability of a financial model, along with evaluating whether it complies with regulations

→ validation methods test machine predictions, what helps to identify problems before deploying the model for real use

→ a conservative method is one that errs on the side of caution, minimizing potential underestimation of risks, losses, or uncertainties;
it is often used in risk-sensitive fields like finance, engineering, and scientific research to ensure safety in drawing conclusions

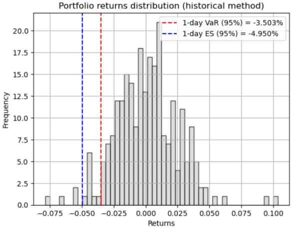
Portfolio construction

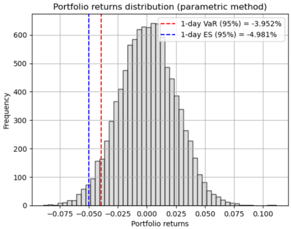
- Tesla, Inc. and the Boeing Company stock data load from Yahoo Finance from within the period 01/03/2024 - 28/02/2025
- assumption of equal stocks' weights in the portfolio (optimization required!)

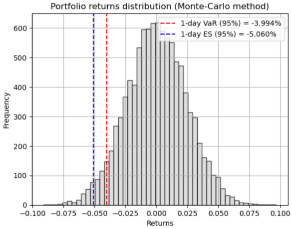
Value at risk (VaR)

→ quantitative measure in the form of a single number summarizing risk of a portfolio
→ denotes a maximum loss in a specified time horizon, with a specified confidence level

The two-asset portfolio VaR for a 1-day time horizon and 95% confidence level using the historical, variance-covariance (parametric) and Monte-Carlo methods:








```
1-day VaR (95%) CALCULATION SUMMARY

 the historical method: -3.503%
 the variance-covariance (parametric) method: -3.952%
 the Monte-Carlo method: -4.011%
```

Backtesting via Kupiec & Christoffersen tests

Kupiec proportion of failures (POF) test (unconditional coverage test)

→ checks if the number of VaR breaches (actual losses exceeding the predicted VaR) matches the expected failure rate
→ if p-value < 0.05 (for 95% cl) → the model fails the test (bad estimate)

Christoffersen conditional coverage (CC) test

→ extends the Kupiec test by checking whether the VaR violations are independent (not clustering in certain periods)
→ if p-value < 0.05 (for 95% cl) → VaR violations are not independent (bad model)

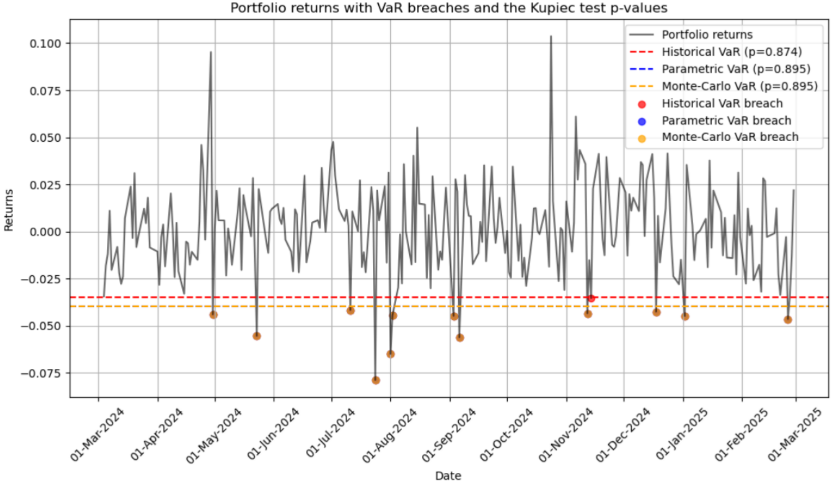
```
 Kupiec's POF test results:
Historical: violations = 13, p-value = 0.8738
Parametric: violations = 12, p-value = 0.8953
Monte-Carlo: violations = 12, p-value = 0.8953

 Christoffersen's CC test results:
Historical: p-value = 0.7014
Parametric: p-value = 0.5964
Monte-Carlo: p-value = 0.5964
```

! test passed, neither model excluded
! historical least conservative

! violations independent, neither model excluded

Backtesting results – visualization



Validation via expected shortfall (ES, C-VaR)

ES denotes expected loss in a specified time horizon, with a specified confidence level, assuming that the loss exceeds VaR.

The two-asset portfolio ES for a 1-day time horizon and 95% confidence level using the historical, variance-covariance (parametric) and Monte-Carlo methods:

```
1-day ES (95%) CALCULATION SUMMARY

 the historical method: -4.95%
 the variance-covariance (parametric) method: -4.981%
 the Monte-Carlo method: -5.216%
```

→ compares the calculated ES with an average of empirical excess losses over the test time period

→ if the ES accuracy ratio differs significantly from 1, the model underestimates risk

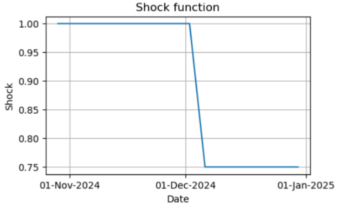
```
 VaR model validation via ES calculations:
Historical: empirical ES = -4.950%, ES accuracy ratio = 1.00000
Parametric: empirical ES = -5.068%, ES accuracy ratio = 1.01755
Monte-Carlo: empirical ES = -5.068%, ES accuracy ratio = 1.00416
```

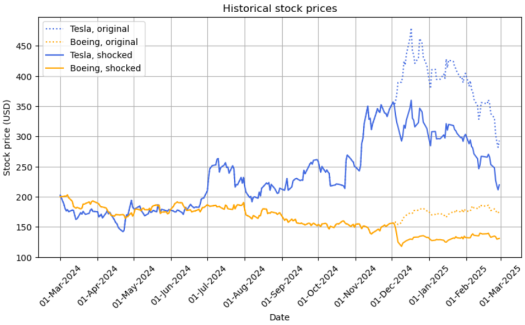
! satisfactory estimates

Stress testing & scenario analysis

→ simulating hypothetical worst-case scenarios for market parameter(s)
→ checking model robustness under extreme conditions

Shock function and stock prices with the market shock of a 25% drop in 4 days applied at the beginning of 12/2024:





Stress testing results

The methods' reaction to extreme losses.

METHOD	historical	parametric	Monte-Carlo
VaR	-3.503%	-3.952%	-4.029%
VaR with shock	-4.372%	-4.221%	-4.191%
Increase in VaR	24.801%	6.826%	4.009%

METHOD	historical	parametric	Monte-Carlo
ES	-4.95%	-4.981%	-5.047%
ES with shock	-5.091%	-5.29%	-5.116%
Increase in ES	2.84%	6.215%	1.364%

! VaR and ES increase after the shock → appropriate response to increased volatility

! VaR and ES losses are not substantial,
! VaR rises more than ES
despite the market shock

} the models might underestimate tail risk, adjustments such as increasing cl to 99% to capture extreme risk better might be needed

Validation outcomes:

- according to the backtesting tests, neither of the VaR estimating models was excluded,
- validation via ES calculations revealed thet the VaR estimates are satisfactory,
- stress testing resulted in conclusion that the models require adjustments to capture extreme risk better.

VaR models validation: broad view

To fully validate a set of VaR models, a combination of methods could be used:

- backtesting for real-world performance comparison, ✓
- expected shortfall (ES, C-VaR) to assess tail risk beyond VaR, ✓
- stress testing to evaluate one-parameter extreme risk scenarios, ✓
- scenario analysis to evaluate extreme complex risk scenarios,
- rolling time windows VaR to check model stability over time,
- distribution tests to validate normality assumptions,
- cross-validation (out-of-sample performance testing, train-test approach) for model generalization assessment.

Thank you for your attention.

Link to the python script containing relevant calculations.