Analyzing VaR estimation accuracy in a portfolio of FAANG stocks using Bayesian Statistics

Gian Marco Oddo - 22-992-952

January 7, 2024

This mini paper presents a Bayesian analysis of Value at Risk (VaR) estimation accuracy in a portfolio of FAANG stocks. The study focuses on examining the precision and reliability of various variance-covariance matrix estimation methods, including the Simple variance-covariance matrix, Linear Shrinkage method, EWMA model, and DCC GARCH model, in calculating the VaR for an equally weighted portfolio over a 504-day period. The Bayesian framework is employed to assess and compare the performance of these methods.

Introduction

FAANG is an acronym that represents some of the most prominent and high-performing technology companies listed on the American stock market: Facebook (now Meta Platforms), Apple, Amazon, Netflix, and Google (now Alphabet). These businesses have been at the pioneering of technological innovation and have had a big impact on investment trends, consumer habits, and advancements in technology worldwide. Their equities are popular for having a large market capitalization, excellent growth potential, and having a significant impact on the direction of the stock market as a whole. Because of their market power, volatility, and position in the tech-driven economy, an analysis of the FAANG stocks is particularly interesting, from a risk perspective. This study examines how they behave together in a portfolio context, specifically assessing the potential effects of various variance-covariance matrix estimation techniques on the computation of the Value at Risk of an equally weighted portfolio formed each day for a two-year period.

Objective of the study

Value at Risk (VaR) is a statistical metric used in finance to quantify the level of financial risk within an investment portfolio over a specific time frame. It represents the maximum potential loss in value of a portfolio over a defined period for a given confidence interval. Essentially, VaR indicates the worst expected loss under normal market conditions within a specific time frame, such as a day or a year, at a certain confidence level.

This study presents a Bayesian analysis of Value at Risk (VaR) for an equally weighted portfolio of FAANG stocks over a period of 504 days. The objective is to examine the precision and reliability of different variance-covariance matrix estimation methods in calculating the daily VaR. The methods employed include the Simple variance-covariance matrix, the Linear Shrinkage method (Ledoit and Wolf), the EWMA model for variances and covariances, and the DCC GARCH model. Each method offers a distinct perspective on capturing the dynamic risk characteristics of the portfolio. In this analysis, the VaR is computed daily using each of the underlying methods for the variance-covariance matrix, based on one year's worth of data. This frequent calibration allows for a detailed tracking of the evolving risk profile of the portfolio.

The formula for Value at Risk (VaR) in the given context is expressed as follows:

$$VaR = \sigma \times q$$

Where:

- σ: The standard deviation (in value) of the portfolio's returns, indicating risk or volatility.
- q: The z-score corresponding to the desired confidence level in a standard normal distribution. For a 95% confidence level, the z-score is approximately 1.645, implying a 95% probability that the portfolio's loss will not exceed the calculated VaR.

Note: This formula assumes a normal distribution of returns.

The calculation of σ (Sigma) is given by:

$$\sigma = \sqrt{x \Sigma x^T}$$

Where:

- x: The vector of value positions in the stocks

Traditionally, it is expected that the actual losses would exceed the calculated VaR on a percentage of days equal to the chosen confidence level α. For example, with a 95% confidence level, the VaR should be exceeded on approximately 5% of the days.
A frequentist would view p as a fixed but unknown parameter and would typically use a Maximum Likelihood Estimation (MLE) method to estimate it.

In contrast, the Bayesian approach treats p as a random variable. This method allows the incorporation of prior knowledge or beliefs about p in the form of a prior distribution. The observed data is then used to update this prior distribution to a posterior distribution, reflecting the updated beliefs about p after considering the evidence.

The Bayesian framework employed here allows for a nuanced and adaptive analysis. It not only tests the accuracy of the VaR calculations against real-world data but also provides insights into the effectiveness of different variance-covariance estimation methods.

Methods used to estimate the Variance Covariance Matrix Σ

1. Simple Historical Method

This method involves calculating the historical covariance matrix.

Formula:

$$\Sigma = \frac{1}{T-1} \sum_{t=1}^{T} (R_t - \bar{R})(R_t - \bar{R})^T$$

Where:

- R_t is the return at time t.
- $\bar{R}$ is the average return over the period.
- T is the number of observations.

2. Linear Shrinkage to the Identity Matrix (Ledoit-Wolf)

This method addresses the instability and estimation error inherent in the sample covariance matrix, particularly when the dimensionality is high relative to the number of observations. By pulling the sample covariance matrix towards the identity matrix (which can be seen as the simplest form of a covariance matrix, where all variables are independent with variance equal to 1), we can improve the estimation's robustness.

Formula:

$$\Sigma_{shrink} = \rho I + (1-\rho)\Sigma_{sample}$$

Where:

- ρ is the shrinkage intensity, which typically lies between 0 and 1.
- I is the identity matrix, representing a simplified prior where all variables are assumed independent.
- Σ_{sample} is the sample covariance matrix obtained from the data.

The shrinkage intensity ρ is a crucial parameter in this method. It determines the balance between the structure imposed by the identity matrix and the structure found in the sample covariance matrix. The goal is to find an optimal ρ that minimizes the mean squared error between the true covariance matrix and the shrinkage estimator. The optimal shrinkage intensity can be estimated using the method proposed by Ledoit and Wolf (2004), which involves the ratio of two quantities:

- The numerator, which measures the squared difference between the sample covariance matrix and the identity matrix, adjusted for degrees of freedom.
- The denominator, which measures the total variation of the sample covariance matrix entries.

3. EWMA Model

The Exponentially Weighted Moving Average (EWMA) model is used for estimating time-varying volatilities and correlations.

Formula:

$$\sigma_{t,i}^2 = \lambda\sigma_{t-1,i}^2 + (1-\lambda)r_{t,i}^2$$
$$\text{cov}_{i,j,t} = \lambda\text{cov}_{i,j,t-1} + (1-\lambda)r_{t,i}r_{t,j}$$

Where:

- λ is the decay factor (0.94).
- $r_{t,i}$ and $r_{t,j}$ are the returns of assets i and j at time t, respectively.
- $\text{cov}_{i,j,t-1}$ is the covariance between assets i and j at time $t-1$.

4. Dynamic Conditional Correlation (DCC) GARCH

The DCC GARCH model is used for modeling time-varying correlations in volatility.

Formula:

$$Q_t = (1-\alpha-\beta)\bar{Q} + \alpha(\epsilon_{t-1}\epsilon_{t-1}^T) + \beta Q_{t-1}$$

Where:

- α, β are parameters estimated from the model.
- ϵ_{t-1} are the residuals from the GARCH model at time $t-1$.
- $\bar{Q}$ is the long-run average correlation matrix.

Components of the Variance-Covariance Matrix (VCM)

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \text{cov}_{12} & \text{cov}_{13} & \cdots & \text{cov}_{1n} \\ \text{cov}_{21} & \sigma_2^2 & \text{cov}_{23} & \cdots & \text{cov}_{2n} \\ \text{cov}_{31} & \text{cov}_{32} & \sigma_3^2 & \cdots & \text{cov}_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{cov}_{n1} & \text{cov}_{n2} & \text{cov}_{n3} & \cdots & \sigma_n^2 \end{pmatrix}$$

Variance σ^2: The diagonal elements of the matrix (σ^2, σ_2^2, ..., σ_n^2) represent the variances of individual asset returns, indicating the degree of fluctuation in each asset's returns.

Covariance (cov): The off-diagonal elements (cov_{ij}) represent the covariances between the returns of different assets, indicating how the returns of one asset move in relation to another.

Results from the Python Code

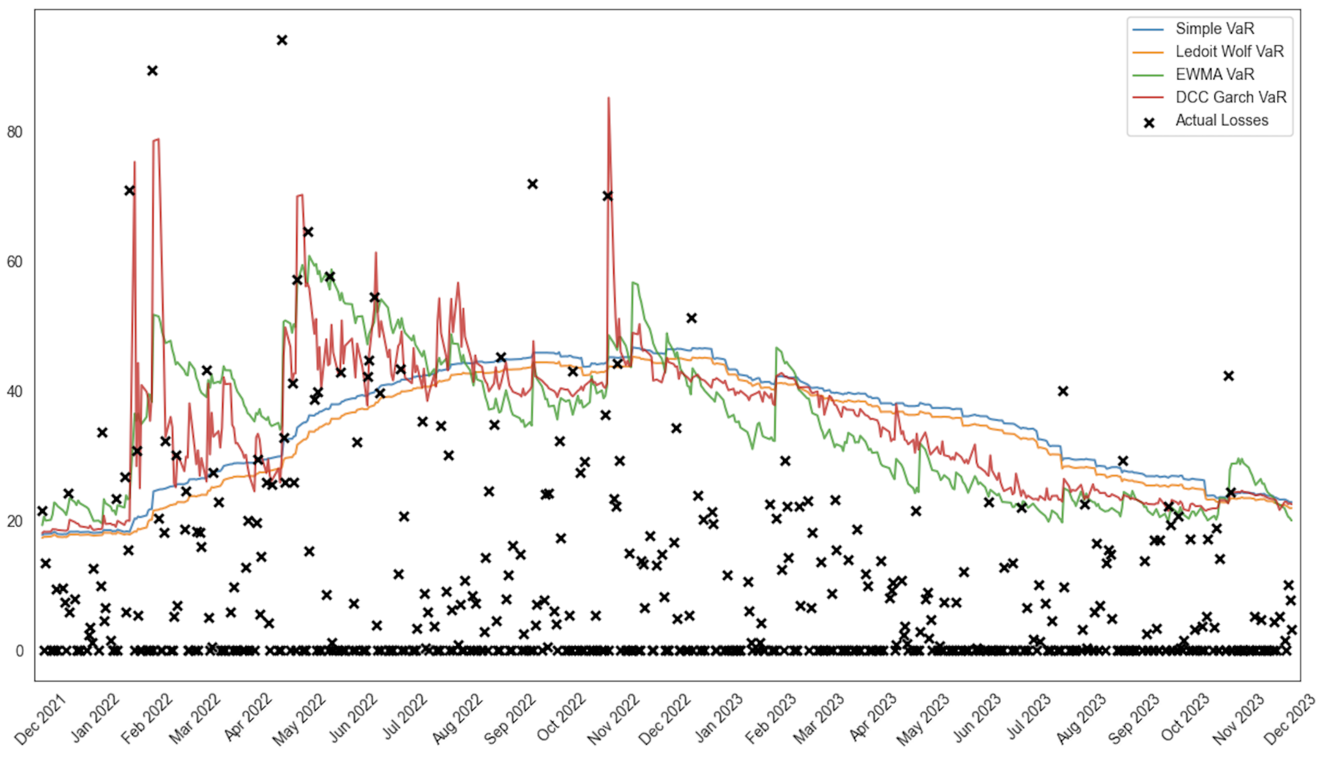

Figure 1: Comparison of various VaR estimation methods, using the four models to estimate the covariance matrix, against actual losses

VaR Model Comparison

The following table compares different Value at Risk (VaR) models based on the number of exceedances:

Model Name	Total Count	Number of Exceedances	Proportion of Exceedances
Simple VaR Exceeded	504	33	6.55%
Ledoit Wolf VaR Exceeded	504	37	7.34%
EWMA VaR Exceeded	504	23	4.56%
DCC Garch VaR Exceeded	504	27	5.36%

Table 1: Comparison of VaRs, derived from the four different models to estimate the covariance matrix, based on exceedances

The selected Prior Distribution for the frequency of VaR exceedances: the Beta Distribution

The Beta Distribution is a versatile and widely used distribution in probability theory and statistics, particularly known for its application in Bayesian analysis. It is defined on the interval [0, 1] and is parameterized by two positive shape parameters, here denoted as α (alpha) and β (beta).

These parameters determine the shape of the distribution, making the Beta Distribution flexible enough to represent a wide range of different probability distributions over a fixed interval.

Probability Density Function (PDF)

The probability density function (PDF) of the Beta Distribution is given by:

$$f(x;\alpha,\beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$$

Where:

- $B(\alpha,\beta)$ is the beta function, which serves as a normalization constant ensuring that the area under the PDF curve is 1.

Expected Value (Mean)

The expected value $E(X)$ of a Beta-distributed random variable X with parameters α and β is given by:

$$E(X) = \frac{\alpha}{\alpha+\beta}$$

Variance

The variance $Var(X)$ of a Beta-distributed random variable X is given by:

$$Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Conjugate Prior Property

One of the key properties of the Beta Distribution in Bayesian analysis is that it is a conjugate prior. This means that if the prior distribution of a probability of success is a Beta Distribution, and the likelihood is binomial (as is common in binary data scenarios), the posterior distribution is also a Beta Distribution. This conjugacy simplifies the Bayesian updating process, making the Beta Distribution particularly useful in sequential data analysis where evidence is accumulated over time.

Given Standard Deviations of the Beta Distributions from an Expert

The following table lists the standard deviations for VaR exceedance as provided by an expert for different variance-covariance estimation methods:

Method	**Standard Deviation**
Simple Historical Variance-Covariance	0.025
Ledoit and Wolf Shrinkage	0.022
EWMA	0.015
DCC GARCH	0.005

Table 2: Expert provided standard deviations for VaR exceedance. The expected value in this case is 0.05 for all the methods

Rationale Behind Different Standard Deviations for VaR Estimation Methods

The specified standard deviations represent expert opinions on the portion of VaR exceedance for different variance-covariance estimation methods. These values are set in relation to alpha, the confidence level of the VaR calculation. The differences in these standard deviations reflect the varying capabilities and assumptions of each method in capturing the risk dynamics of a financial portfolio.

1. **Simple Historical Variance-Covariance Matrix (0.025)**: This method, with a standard deviation of 0.025, is a straightforward approach but does not account for volatility clustering. It assumes that historical volatilities and correlations are stable over time, which might not accurately reflect market realities.

2. **Ledoit and Wolf Shrinkage Method (0.022)**: This approach, with a slightly lower standard deviation of 0.022, improves upon the simple historical method by shrinking the extreme values in the variance-covariance matrix towards a central estimate. However, it still does not explicitly address volatility clustering.

3. **EWMA Model (0.015)**: The Exponentially Weighted Moving Average (EWMA) model, with a standard deviation of 0.015, accounts for volatility clustering by giving more weight to recent observations. This helps in capturing the changing volatility patterns but still operates under the assumption of a normal distribution for returns.

4. **DCC GARCH Model (0.005)**: The Dynamic Conditional Correlation (DCC) GARCH model, with the lowest standard deviation of 0.005, is the most sophisticated among the four. It not only addresses volatility clustering but also models the long-term correlations between different assets.

The variance in standard deviations underscores the importance of model selection in VaR estimation. Models that account for volatility clustering and long-term correlations (like the EWMA and DCC GARCH) are perceived to be more accurate, as indicated by their lower standard deviations in the expert's assessment. However, it's important to mention that all these methods assume normality in the VaR calculation formula, which could be a limitation if the actual return distributions deviate from normality.

These expert opinions, combined with the expected values of their respective Beta Distributions, provide the foundational parameters (alpha and beta) for the initial prior distributions in the Bayesian analysis.

This approach is crucial in establishing a well-informed starting point for the Bayesian inference process, allowing for a more nuanced understanding and prediction of VaR exceedances in the context of high-volatility portfolios such as those comprised of FAANG stocks.

Prior Densities Derived from Experts' Opinions

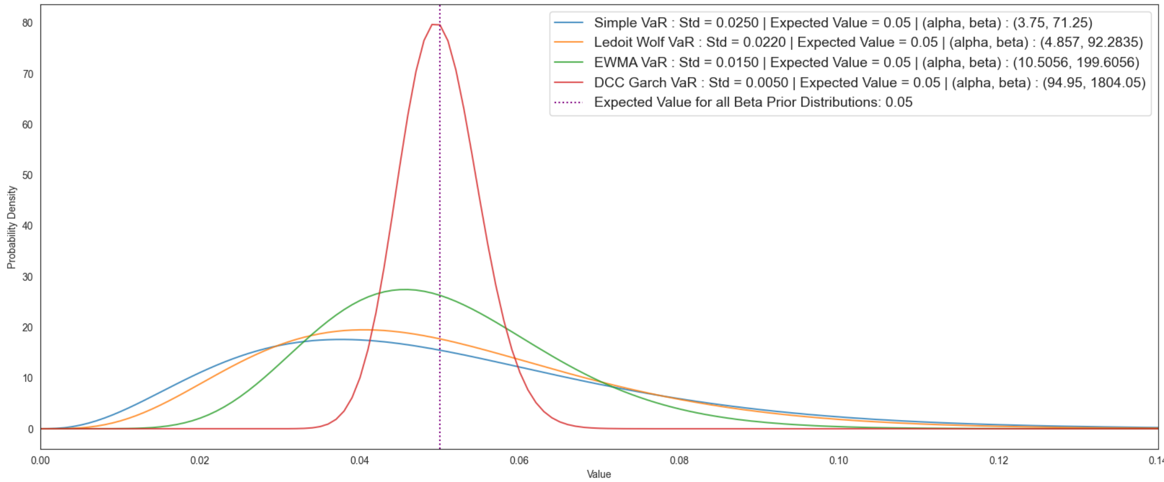

Figure 2: Prior density functions for the four Beta Distributions that reflect expert opinions

Posterior Distribution for the % of times VaR is exceeded: the Core of the Bayesian Approach

In every Bayesian analysis, the process of going from the prior distribution to the posterior distribution involves updating our beliefs about the parameters of interest after observing new data. This is done using Bayes' Theorem, which in the context of parameter estimation is typically written as:

$$\text{Posterior}(\theta|data) \propto \text{Likelihood}(data|\theta) \times \text{Prior}(\theta)$$

Where:

- $\text{Posterior}(\theta|data)$ is the probability distribution of the parameter θ after observing the data.
- $\text{Likelihood}(data|\theta)$ is the probability of the observed data under a given parameter value θ.
- $\text{Prior}(\theta)$ is the probability distribution of the parameter θ before observing the data.
- $\propto$ denotes proportionality; to get the actual posterior distribution, we need to normalize this product so that it integrates to 1.

For the Beta Distribution as a prior in the case of binomial data (such as VaR exceedances), the likelihood function for observing x successes out of n trials is given by the binomial distribution:

$$\text{Likelihood}(x|n,p) = \binom{n}{x} p^x (1-p)^{n-x}$$

Here, p represents the probability of success on each trial, which in the context of VaR would be the probability of exceedance.
If the prior distribution for p is $\text{Beta}(\alpha, \beta)$, the posterior distribution after observing x successes and $n-x$ failures is also a Beta Distribution, as previously stated, specifically:

$$\text{Posterior}(p|x,n) = \text{Beta}(\alpha + x, \beta + n - x)$$

This reflects the conjugate prior property of the Beta Distribution: the posterior parameters are simply the prior parameters updated by the number of successes and failures.

Expected Value and Variance of the Posterior Distributions

The expected value $E(p)$ of the posterior Beta Distribution is:

$$E(p) = \frac{\alpha + x}{\alpha + \beta + n}$$

The variance $\text{Var}(p)$ of the posterior Beta Distribution is:

$$\text{Var}(p) = \frac{(\alpha + x)(\beta + n - x)}{(\alpha + \beta + n)^2(\alpha + \beta + n + 1)}$$

Credible Intervals vs Confidence Intervals

In Bayesian analysis, a credible interval provides the range within which an unobserved parameter value falls with a particular subjective probability. This is in contrast to confidence intervals used in frequentist statistics, which provide a range that would contain the parameter with a certain frequency if the experiment were repeated many times.

Credible intervals are derived directly from the posterior distribution and have a clear interpretation: the "95% credible interval" means there is a 95% probability that the parameter lies within this interval, given the observed data and the prior information.

On the other hand, a "95% confidence interval" in the frequentist context means that if we were to repeat the experiment many times and calculate the confidence interval each time, 95% of those intervals would contain the true parameter value. It does not assign a probability to the parameter being within a specific interval from one experiment.

This difference in interpretation stems from the fundamental distinction between Bayesian and frequentist approaches: Bayesian inference treats parameters as random variables with associated probability distributions, while frequentist inference treats parameters as fixed but unknown quantities.

To calculate a credible interval for a Beta Distribution:

1. Determine the posterior Beta Distribution parameters based on prior parameters and observed data.
2. Use the cumulative distribution function (CDF) of the Beta Distribution to find the bounds of the interval. For a 95% credible interval, we find values a and b such that $P(X \leq a) = 0.025$ and $P(X \leq b) = 0.975$, where X follows the Beta Distribution.
3. The interval $[a, b]$ is the 95% credible interval, indicating, as previously stated, that there is a 95% probability that the parameter lies within this interval given the observed data and prior.

In general, the cumulative distribution function (CDF) of a Beta Distribution for a random variable X with parameters α and β is given by the integral form:

$$CDF(x; \alpha, \beta) = \int_0^x \frac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha, \beta)} dt$$

where $B(\alpha, \beta)$ is the Beta function. The values a and b for the 95% credible interval are found by solving for x in the equations $CDF(a; \alpha, \beta) = 0.025$ and $CDF(b; \alpha, \beta) = 0.975$.

Posterior Beta Distributions for the four methods

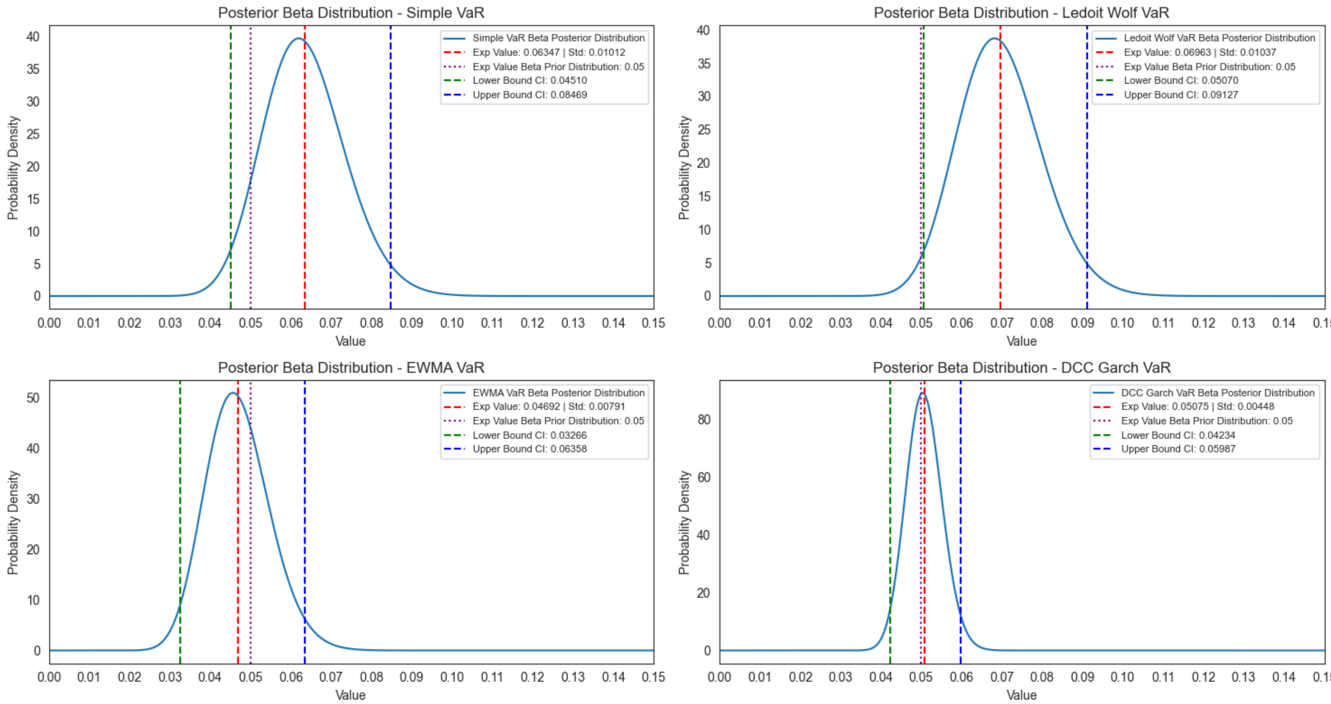

Figure 3: Posteriors density functions with expected value and 95% credible interval, after the influence of the data

Results

Method	**Exp. Value**	**SD**	**LB CI**	**UP CI**
Simple VaR	0.06347	0.01012	0.04510	0.08469
Ledoit Wolf VaR	0.06963	0.01037	0.05070	0.09127
EWMA VaR	0.04692	0.00791	0.03266	0.06358
DCC Garch VaR	0.05075	0.00448	0.04234	0.05987

Table 3: Posterior expected values, standard deviations and credible interval for the four methods

The analysis yields a definitive outcome: the DCC GARCH model emerges as the superior method for estimating the Value at Risk (VaR). This conclusion is supported not only by prior beliefs, which favor the DCC GARCH model for its proficiency in capturing dynamic correlations and because it represents a weighted average of conditional correlation and last term trend, but also by empirical evidence. Among the methods considered, the DCC GARCH model demonstrated the lowest variance, aligning with expert expectations and reinforcing its reliability as confirmed by the data: its posterior expected value for the frequency of VaR exceedances is notably aligned with the VaR's alpha level.

Furthermore, the credible interval obtained from the DCC GARCH model is the most precise compared to the others, offering tighter and more confident range of the estimated parameter. Additionally, the standard deviation of the posterior distribution is the smallest of all four methods examined, indicating the highest level of precision in the estimates provided by the DCC GARCH model.

Prior VS Posterior Beta Distributions

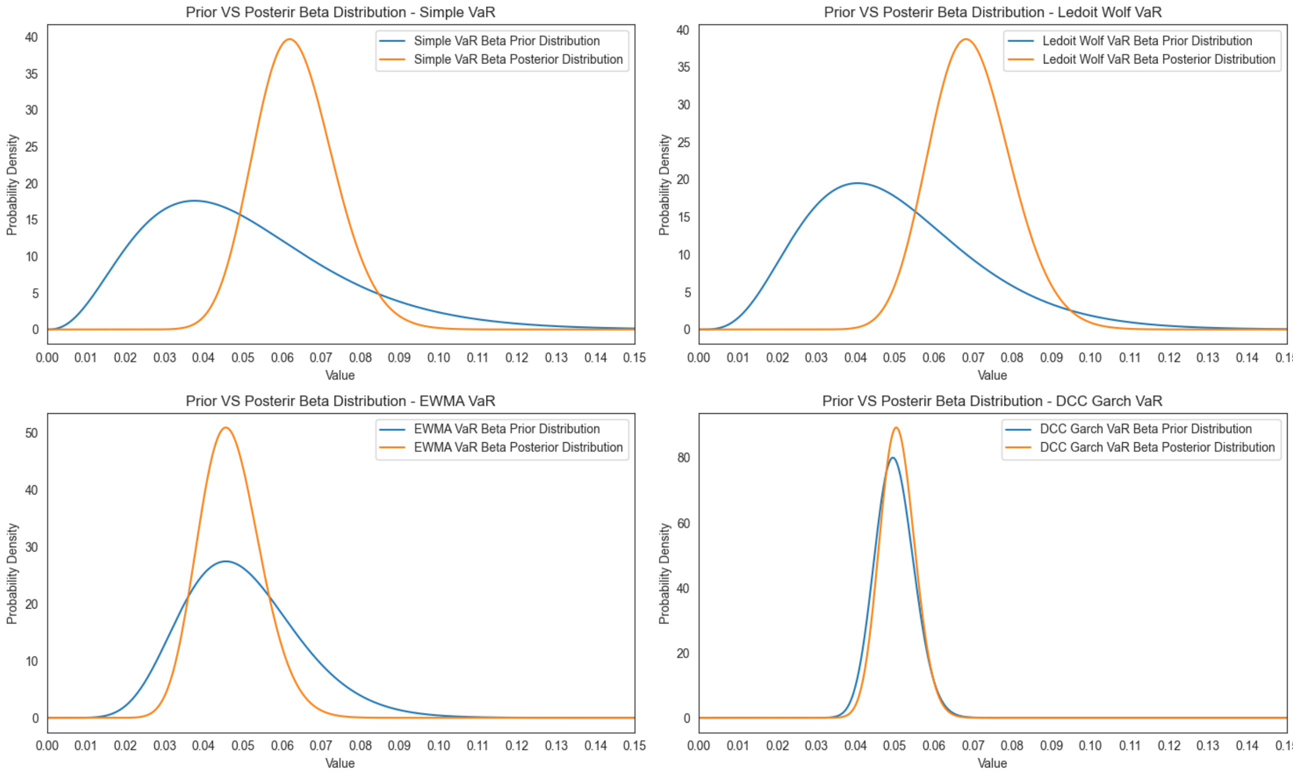

Figure 4: Prior vs posterior density functions for the Beta Distributions, after the influence of the data

Simple VaR

The prior distribution's peak is notably lower than the posterior's, indicating that the observed data suggest a higher probability of VaR exceedances than initially assumed. This considerable shift may imply that the Simple VaR method's initial model assumptions were conservative, or it did not fully capture the risk inherent in the data.

Ledoit Wolf VaR

The posterior distribution shifts to the right here as well, echoing the Simple VaR's findings but with a more pronounced movement. This suggests that the data did lead to a higher estimated probability of VaR exceedances.

EWMA VaR

The posterior shift is minimal to the left, which could indicate that the EWMA VaR's prior assumptions were more in line with the observed data. A lower posterior variance indicates that after accounting for the new data, there is increased certainty in the model's predictions. This increased precision implies that the model has a high degree of confidence in the estimated probability of VaR exceedances and suggests that the risk estimates provided by the EWMA VaR are reliable

DCC Garch VaR

The posterior distribution's close alignment with the prior suggests that the data closely matched the initial model expectations, which may indicate a strong fit of the DCC Garch model. This minimal shift signifies that the model, with its dynamic nature, was likely accurate in capturing the underlying risk from the start.

Closing Remarks

A shift in the peak from the prior to the posterior distribution represents an update in belief about the probability of VaR exceedances after considering new data. The extent of this shift is telling of how the new data affect prior beliefs:

- A large shift might indicate that the original model was not fully capturing the risk, prompting a significant update in the probability assessment.
- A minimal shift suggests that the observed data were well anticipated by the model, which might not necessitate substantial changes to the risk estimation.

The DCC GARCH VaR model, among the four tested, stands out for its remarkable reliability. Its close alignment with the 5% alpha level demonstrates that the observed data further supports the initial expert opinion. This convergence, achieved through a Bayesian approach, makes it particularly effective for estimating the Value at Risk.

Reference

1. Olivier Ledoit and Michael Wolf, "Honey, I Shrunk the Sample Covariance Matrix." Equities Division, Credit Suisse First Boston, London, UK, and Department of Economics and Business, Universitat Pompeu Fabra, Barcelona, Spain. November 2003.
2. Elisabeth Orskaug, "Multivariate DCC-GARCH Model - With Various Error Distributions." Master of Science Thesis, Department of Mathematical Sciences, Norwegian University of Science and Technology. Supervisor: Håvard Rue, MATH; Co-supervisor: Kjersti Aas, Norsk Regnesentral. June 2009.
3. Axel A. Araneda, "Asset volatility forecasting: The optimal decay parameter in the EWMA model." Institute of Financial Complex Systems, Faculty of Economics and Administration, Masaryk University, Brno, Czech Republic.
4. Weiqian Li, "Value at Risk (VaR) and its calculations: an overview." Instructor: Wenqing Hu, MATH 4997W, Date: 04/25/2016.