FIM-549
Financial Risk Analysis

Project Report

Submitted By-

Siddharth Thakur

NCSU Master - Financial Mathematics ‘2022

Introduction

Scope of the Project

This project aims to provide an insight into the real time experience for quantitative modelling for finance. The structure of the project follows the lifecycle of a real time project in a financial institution looking to emulate workplace tasks, responsibilities, and deliverables. The project should be looked in its entirety with the accompanying code and data files, however, this report aims to provide a thorough overview of the entire project.

This project involves choosing a company (JP Morgan Co.) and analyze the historical market data for the company against an index and evaluate risk measures form the perspective of the company. This is done in two broad parts. The first part involves the use of historical data to forecast volatilities using a GARCH model and analyzing the changing volatilities for the market price of the company. This serves as a base for analyzing the risk for the company in comparison to an index.

The second part of the project focuses on analyzing the risk measures of three portfolios created from the company shares and the index. The scope here is limited to estimation of Value at Risk, Expected shortfall for these portfolios and back testing for it. The advantages and limitation of the methods are also discussed.

Being an academic project, this report also answers few of the question that were part of the problem statement and provides evidence when needed.

The project material consists of the following deliverables:

- Project Report
- Jupyter Notebook for Python Code - Code.ipynb
- Excel CSV for data - Adjusted_close_Data.csv
- Excel Sheet used to create tables - Tables and Information.xlsx

Structure of the Project

Data Preparation

The data in this project is sourced from Yahoo Finance[7]. The "Adjusted Value" is extracted using the "Pandas_DataReader" Library in Python. The following function from the library is used to get the final dataset from the Yahoo Finance website.

```
DataReader([TICKER], [SOURCE], start=[Start Date], end=[End Date])
```

The inputs were varied accordingly.

The data for the following Tickers is extracted using the API. The following are the details for data extracted.

Tickers	Description	Data Used	Start Date	End Date
JPM	JPMorgan Chase & Co	Adjusted Closing price	January 1, 2019	February 28, 2022
SPY	SPDR S&P 500 ETF Trust	Adjusted Closing price	January 1, 2019	February 28, 2022

The final data set from the source can be found here: [2]

Finally, using the above data set, the code calculates the log percentage returns for each day for each of the tickers. The formula used here is

$$Log_{Returns} = \frac{Log\left(\mathrm{Pr}_{t}\right) - Log\left(\mathrm{Pr}_{t-1}\right)}{Log\left(\mathrm{Pr}_{t-1}\right)} * SCALAR$$

$$\mathrm{Pr}\ \ is\ the\ Adjusted\ Closing\ price$$

$$SCALAR = 100\ in\ this\ case\ to\ accomodate\ modelling\ needs$$

For further processing the data was split by date. The details for the same can be found in the relevant sections.

Part 1

Volatility Modelling Using GARCH

The problem statement requires fitting the GARCH (1,1) model to the log returns to estimate the volatility for the log returns of both the tickers. The data used for this for this exercise is ranging from January 1, 2019, to December 31, 2021. This section will describe the theoretical basis and tasks accomplished in thei project towards a correct implementation.

GARCH Model - Introduction

The GARCH[4][5] model was proposed by Bollerslev in 1986 as an extension to the EWMA (Exponentially Weighted Moving Average) model. The EWMA model looks to model the volatility as a function of its historical values. The model assigns unequal weights to the historical values. The EWMA model specifically assigns decreasing weights to past values thereby giving more importance to later values. The formula used in the EWMA model is given below. Here the λ denotes the weight and is a parameter to the model that needs to be estimated to the data.

$$\sigma_n^2 = \lambda\sigma_{n-1}^2 + (1-\lambda)u_{n-1}^2$$

The GARCH model assumes there is clustering in periods of high or low volatility, so it aims to exploit the recent volatility to predict volatility in the near future. The more general GARCH (p, q) model calculates σ^2 n from the most recent *P* observations on μ^2 and the most recent *Q* estimates of the variance rate. The general equation for the GARCH (P, Q) model is :

$$a_t = \sigma_t \varepsilon_t, \qquad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2$$

In this context, the GARCH model used has P and Q set to 1. Therefore, a GARCH (1,1) model is used in modelling volatility from a long-run average variance rate, VL, as well as from σ_{n-1} and u_{n-1} (last change observed in the data). The equation used is:

$$\sigma_n^2 = \gamma V_L + \alpha u_{n-1}^2 + \beta\sigma_{n-1}^2$$

The limiting condition here is that the total weights assigned should sum to 1. Hence

$$\gamma + \alpha + \beta = 1$$

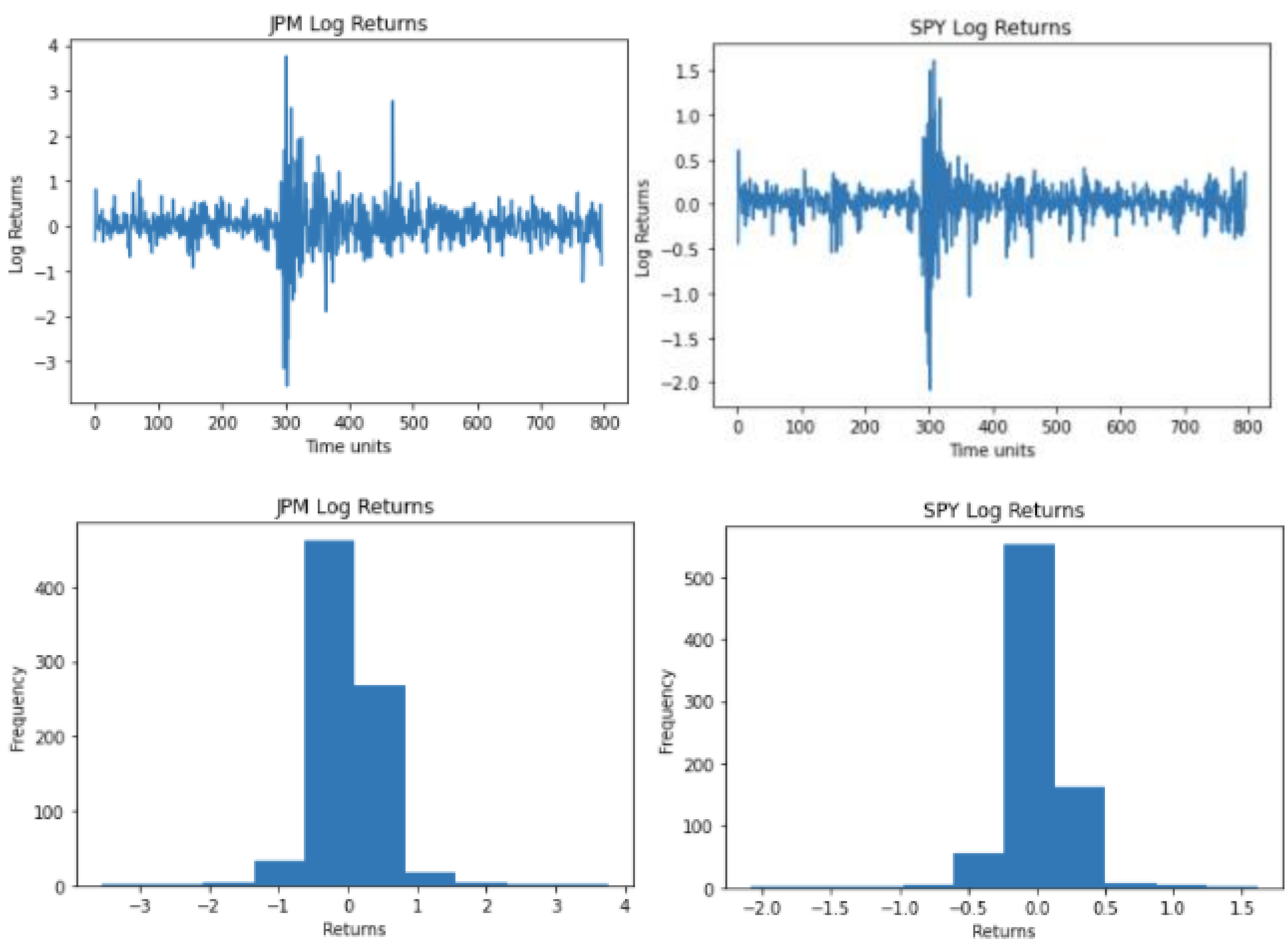
The assumptions of the model are:

- The variance of the error term follows an autoregressive moving average process and the process in general has a stationary mean (usually 0).
- The returns series is conditionally normally distributed. This is referred to this as the normal GARCH model

GARCH (1,1) Modelling

Assumption Check

The Log returns were plotted to check for normality. The graphs are shown here:



The histogram show that we can assume the data to be (conditionally) normally distributed. The returns plotted above show what we mean by varying volatility with time. The ‘heteroscedasticity” of the date can be seen well from the plots as we see period with low volatility and high volatility.

Model Fitting

This section serves as answer to questions 2.1 – 2.2

The data for the volatility was fit using the ARCH package in python. The model uses inputs form the user to estimate parameter for the GARCH model using the 'Maximum Likelihood Method'. However, the package requires the values in the data to be greater than 1 since using a smaller scale of value hinders in accurate estimation of the parameters. Therefore, the returns were scaled using a factor of 100. This was referenced in the data preparation section.

The data used to train the model ranged from January 1, 2019, to December 31, 2021 as per the requirement in the problem statement.

The model results are as follows:

GARCH MODEL RESULTS FOR JPM		
Fit Result		
Estimate	**Coefficient**	**P - Value**
Omega	0.0060045	0.011830
Alpha[1]	0.1896	0.002340
Beta[1]	0.7774	0.000000
Estimated		
Estimate	**Coefficient**	**Formula**
Gamma	0.033	1 - Alpha - Beta
V_L	0.004266	Sqrt(Omega / (Gamma * Scalar ^2))

GARCH MODEL RESULTS FOR SPY		
Fit Result		
Estimate	**Coefficient**	**P - Value**
Omega	0.0018954	0.001562
Alpha[1]	0.2578	0.000011
Beta[1]	0.7045	0.000000
Estimated		
Estimate	**Coefficient**	**Formula**
Gamma	0.0377	1 - Alpha - Beta
V_L	0.0022360	Sqrt(Omega / (Gamma * Scalar ^2))

The model estimates are all significant based on the P values of the estimates. The values of Gamma and V_L (Long term Volatility) are calculated from the model estimates as per the specifications in the previous section.

Since the original data was scaled, the Omega term needs to be scaled back to get the correct value. The correct scalar to apply to Omega is $Scalar^2$ since this is an intercept in a quadratic equation[1].

Model Forecast

This section serves as answer to question 2.3

The forecast for the model was done using the GARCH Forecast method within the ARCH package. Internally, the forecast method used the formula for the GARCH model. The forecasts of the volatility for 3rd January 2022 are:

These estimates have been adjusted for the scalar.

Volatility Forecast	
Ticker	**January 3, 2022**
JPM	0.020554562
SPY	0.01239274

Correlation Estimation

This section serves as answer to question 3

The correlation between the returns of JPM and SPY have been calculated using the Pearson product-moment correlation formula.

$$R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} * C_{jj}}}$$

Here:

C_{ij} denotes the covariance between item i and j (equals variance for i=j).

R_{ij} denotes the correlation coefficient as needed.

The correlation matrix estimated for the full range of data (Jan 1, 2019 to Feb 28, 2022) is given below:

Correlation Coefficient Matrix		
Tickers	**JPM**	**SPY**
JPM	1	0.7648478
SPY	0.7648478	1

The data shows a fairly large (positive) correlation between the prices of the two tickers. This is as expected since JP Morgan Chase Co. is a very large company and is part of the S&P 500 Index.

Part 2

Portfolio Creation

For this section, there are three portfolios created and invested in as of December 31st, 2021. The 3 portfolios are:

A. $1 million invested in JP Morgan Chase co. (JPM)
B. $1 million invested in S&P 500 Index ETF (SPY)
C. $2 million invested in both the above stocks with equal weights.

Value at Risk (VaR) and Expected Shortfall (ES)

Value at Risk(VaR)[6] is a statistic that quantifies the extent of possible financial losses within a firm, portfolio, or position over a specific time frame. This metric is most commonly used by investment and commercial banks to determine the extent and probabilities of potential losses in their institutional portfolios.

Risk managers use VaR to measure and control the level of risk exposure. One can apply VaR calculations to specific positions or whole portfolios or use them to measure firm-wide risk exposure.

The value at risk and expected shortfall were calculated using the historical simulations. It is one method to calculate the VaR and Expected Shortfall.

Historical Simulations - Methodology

The methodology followed in this approach is as follows.

- Assume that the VaR for a portfolio needs to be calculated using a one-day time horizon, a 99% confidence level, and 501 days of data.
- The first step is to identify the market variables affecting the portfolio. These market variables are sometimes referred to as risk factors. They typically include exchange rates, interest rates, stock indices, volatilities, and so on.
- Data are then collected on movements in these market variables over the most recent 501 days. This provides 500 alternative scenarios for what can happen between today and tomorrow.
- For each scenario, the dollar changes in the value of the portfolio between today and tomorrow is calculated. This defines a probability distribution for daily loss (with gains counted as negative losses) in the value of the portfolio.
- The estimate of VaR is the loss at the 99th- percentile point. ES can be estimated by averaging the losses that are worse than VaR.

Estimation

This section serves as answer to question 4.1

The estimation of the data is done using the same methodology. The training data was based on the dates from Jan 1, 2019, to Dec 31, 2021. This created 757 simulations of returns which were used in the calculations. The losses for all these simulations were sorted in descending order and the 1^{st} percentile was selected. In this case the 1^{st} percentile came to be the 7.57^{th} value. To account for the fractional value, the weighted average of the 7^{th} and 8^{th} values was calculated. Hence

$$99\%\ VaR = (0.57 * 7^{th} Value)\ + (0.43 *\ 8^{th}\ Value)$$

The Expected Shortfall (ES) is calculated as the average of all the values worse than the VaR. Hence, the value of the 99% ES is calculated as the average of the 7 (since the 1 percentile is 7.57) greatest losses.

The result estimates are given as below:

Portfolio	Description	Value at Risk (VaR)	Expected Shortfall (ES)
A	**$1 million- JPM**	$ 62,327.30	$ 104,580.93
B	**$1 million- SPY**	$ 44,964.90	$ 73,370.99
	Total (A+B)	*$ 107,292.20*	
	Diversification benefit	*$ (8,505.85)*	
C	**$2 million- JPM & SPY**	$ 98,786.36	$ 175,951.26

The diversification benefit here is calculated as the difference between the sum of the individual VaR of the portfolio A & B (since they are created using the same components) and the VaR of portfolio C.

Back Testing

Introduction

Back-testing[4] is an important reality check for a risk measure. It is a test of how well the current procedure for calculating the measure would have worked in the past. Assuming a 99% VaR needs to be tested, Back-testing involves looking at how often the loss in a day would have exceeded the one-day 99% VaR when the latter is calculated using the current procedure. Days when the actual loss exceeds VaR are referred to as exceptions. If exceptions happen on about 1% of the days, we can feel reasonably comfortable with the current methodology for calculating VaR. The period of the test should be the previous 250 days as prescribed in the BIS 98 method.

A similar approach was taken to back test the estimated VaR on the data ranging from March 1st, 2021- February 28th, 2022, for portfolio A (As required by problem statement). For this purpose, scale of the portfolio used was $1 million.

Test Results

This section serves as answer to questions 4.2- 4.3

Back Test Results	
Exceptions	0

The back test of the 99% VaR for portfolio A (JPM) gave very interesting results. There were 0 exceptions in the back test. This means that the multiplier to be set to be 3 according to the BIS 98 (Basel 1996 amendment). Therefore, the capital required for the portfolios would be

$$\max(\mathrm{VaR}_{t-1}, m_c \times \mathrm{VaR}_{\mathrm{avg}}) + \mathrm{SRC}$$

Where SRC is the specific risk charge and m_c is the multiplicative factor set.

This is very surprising since we expect around 2-3 exceptions to occur since the length of the data being tested in 252 and the testing is being done using a VaR at 99% confidence interval.

This result can be interpreted as follows:

- The 99% VaR calculation may be overestimating the actual Value at Risk at 99% Confidence Interval.
- This might be an instance of Bunching. Since we have used a subset of the simulation data, the date range chosen for the back test does not contain any values worse than the VaR estimate. The report discusses bunching and how to adjust for it in the next section.
- The VaR is estimated using the fact that the returns are normally distributed, but this might not be the case. In reality, extreme results have a heavier tail than the normal distribution. We might use Extreme value theory to estimate the VaR. This would estimate the distribution of extreme results from extreme data only and avoid the excess noise of mid-section data.

Bunching

This section serves as answer to question 4.4

Bunching[4] is the reason we have abnormal results in the back test. If daily portfolio changes are independent, exceptions should be spread evenly throughout the period used for back testing. In practice, they are often bunched together, suggesting that losses on successive days are not independent.

One way to test for Bunching is to use the statistic suggested by Christoffersen [3] as follows:

$$-2\ln[(1-\pi)^{u_{00}+u_{10}}\pi^{u_{01}+u_{11}}] + 2\ln[(1-\pi_{01})^{u_{00}}\pi_{01}^{u_{01}}(1-\pi_{11})^{u_{10}}\pi_{11}^{u_{11}}]$$

$$\pi = \frac{u_{01}+u_{11}}{u_{00}+u_{01}+u_{10}+u_{11}}$$
$$\pi_{01} = \frac{u_{01}}{u_{00}+u_{01}}$$
$$\pi_{11} = \frac{u_{11}}{u_{10}+u_{11}}$$

where u_{ij} is the number of observations in which we go from a day where we are in state i to a day where we are in state j. This statistic is chi-square with one degree of freedom if there is no bunching.

Extreme Value Theory

This section serves as answer to question 4.4

Extreme value theory (EVT)[4] is the term used to describe the science of estimating the tails of a distribution. EVT can be used to improve VaR or ES estimates and to help in situations where analysts want to estimate VaR with a very high confidence level. It is a way of smoothing and extrapolating the tails of an empirical distribution.

This can be used to get a smoother, more accurate distribution of extreme losses.

Normal Approximation to VaR and ES estimation

This section serves as answer to questions 5.1-5.2

This method assumes that the returns of the portfolios are normally distributed. Hence the empirical estimates of the population mean, and variance can help create a confidence level for estimating VaR.

Results	
Investment	$ 2,000,000.00
Correlation(A,B)	0.7648478
Weight A	0.5
Weight B	0.5

Portfolio	Mean	Standard Deviation
A	0.00	0.022186
B	0.00	0.013761
C	0.00	0.016945

Z0.01		-2.326347874
VaR (per unit)		(0.04)
VaR ($2 million)		$ 78,841.86

Here the following formulas have been used to estimate the VaR for portfolio C.

$$Var(C) = a_1 Var(A) + a_2 Var(B) + 2 * Corr(A,B) * a_1 * a_2 * \sqrt{(Var(A) * Var(B))}$$

$$\text{Value at risk VaR(C)} = \text{Mean(C)} + \sqrt{Var(C)} * Z_{0.01}$$

Where $Z_{0.01}$ is the 0.01 percentile for standard normal distribution which is = -2.3263 as above. Since VaR is a loss, the table does not show the negative sign.

The VaR for portfolio C calculated using historical simulations was $98786.35 but the normal approximation estimates it to be around $78,841.86. **The normal approximation underestimates the VaR for the portfolio.**

This is as expected since a normal distribution has very thin tail compared to the real time experience for stock returns. Other model does not consider light tails and hence give a larger estimate of the VaR.

RAROC Estimation

This section serves as answer to questions 6

Risk-adjusted performance measurement[4] (RAPM) has become an important part of how business units are assessed. There are many different approaches, but all have one thing in common. They compare return with capital employed in a way that incorporates an adjustment for risk.

The most common approach is to compare expected return with economic capital (VaR - Expected losses). This is usually referred to as RAROC (risk-adjusted return on capital). The formula is:

$$\text{RAROC} = \frac{\text{Revenues} - \text{Costs} - \text{Expected losses}}{\text{Economic capital}}$$

The approach followed here is like the one mentioned above. We calculate the economic capital which is assumed to be equal to the estimated VaR for the portfolio. Assuming no costs and expected losses we calculate the returns for the portfolios which is the percentage change in the stock prices multiplied by the investment amount.

The price for Portfolio C is calculated as the weighted average of the returns of portfolio A and B. The weights used are the amounts invested and total price of the portfolio. The full calculations can be referenced in the Excel calculation sheet.

Portfolio	Economic Capital	Price 12/31/21	Price 1/3/22	Returns	Profit/Loss	RAROC
A	$ 62,327.30	156.24	159.55	0.021185	$ 21,185.36	**33.990%**
B	$ 44,964.90	473.48	476.23	0.005808	$ 5,808.06	**12.917%**
C	$ 98,786.36	314.86	317.89	0.009623	$ 19,246.65	**19.483%**

The above results show that portfolio A gives the highest risk adjusted results. This is plausible because JPM has a good return on that day. However, we should look at the RAROC for a longer period of time to correctly assess the performance.

Conclusion

The following were the observations and lessons learnt form the project and results.

- The theory pertains to very ideal situations which does not exist is real life.
- Risk cannot be looked at in isolation. A wholistic view of the situation is required.
- VaR and Expected Shortfall are only a small part of the risk management process, there are other factors and attributes that need to be considered when dealing with market risk.

Appendix

- [1] - https://stats.stackexchange.com/questions/380174/do-parameters-stay-unchanged-when-garch-is-scaled
- [2] – Data Link
- [3] - See P. F. Christoffersen, “Evaluating Interval Forecasts,” International Economic Review 39 (1998): 841–862
- [4] - Risk Management and Financial Institutions – John C Hull
- [5]- GARCH Process (investopedia.com)
- [6] - Value at Risk (VaR) Definition (investopedia.com)
- [7] - Yahoo Finance - Stock Market Live, Quotes, Business & Finance News