WORCESTER POLYTECHNIC INSTITUTE

MARKET AND CREDIT RISK MODELS AND MANAGEMENT

Risk Management Project

Author:
Yijiang XU
Weixi LIU

Instructor:
Professor Marcel Blais

April 30, 2018

Contents

List of Figures

Chapter 1

Introduction

In this project part 1, we first construct a portfolio with 15 stocks traded on American exchanges with equal weight, and calibrate three different type of models for unconditional loss distributions from our portfolio's monthly loss series, where two models based on normal distribution and the another one based on t-distribution. we do risk analysis with our portfolio's monthly loss among the three models by using value at risk and expected shortfall risk measures, and also do comparisons among the three models to check which one's expected monthly loss is closer to our actual monthly losses.

In part 2, it's mainly about the estimation of value at risk and expected shortfall using polynomial tails. We use the semiparametric model for our estimation and construct a historical time series of returns of our portfolio in order to get the proper parameters a and A. We compare the difference results of a between the Hill estimator method and regresssion estimator method.

In part 3, we mainly focus on the ARMA(1,1)-GARCH(1,1) model for our conditional loss distribution to calculate the VaR and ES. After we get the results, we compare the conditional risk management with the unconditional risk management and also compare the risk analysis with the previous ones.

In final part, we perform the risk management with our own holding portfolio. First, we use the loss approximations to estimate the loss distribution of our portfolio with options and bonds. Then we perform a SVaR-based Stress Test for our portfolio. After doing that, we try to perform a 15 percent risk reduction with our portfolio by buying put option. Finally, using the previous methods we perform a risk analysis with our own holding portfolio.

Chapter 2

Data Collection

Through Yahoo Finance[1], we download 6-years AAL, AAPL, AEO, AMZN, CAR, DLTR, GOOG, KNX, MCD, NFLX, NKE, NCDA, SINA, TSLA, and UPS historical data from 01/03/12 to 12/29/17. Through Bloomberg, we download the data (such as price table and Greeks) we need for KNX US 05/18/18 C50 Equity[2], NFLX US 05/18/18 C315 Equity, XEO UO 05/18/18 P1220 Index, and two US Treasury Bonds from 02/26/18 to 04/19/18 which is the holding period of our portfolio. Python Pandas package helps us to read excel file into data frame. For stock, we extracted the adjust closed prices separately and transferred them into 1-dimensional arrays. They would be accessed in all problems of the project.

[1] `https://finance.yahoo.com/`

[2] It means that a KNX call option with strike at 50 and maturity is 05/18/18.

Chapter 3

Methodology

3.1 Returns

Def 1. *The new return over holding period [t-1,t] is (Blais, 2017)*

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}.$$

Def 2. *The (simple) gross return is (Blais, 2017)*

$$\frac{P_t}{P_{t-1}} = 1 + R_t.$$

Def 3. *The log returns are defined as (Blais, 2017)*

$$r_t = \ln(1 + R_t).$$

3.2 Loss distribution

Assume V_t is our portfolio price on day t, then we can define a loss of portfolio in this way:

Def 4. *For a given time horizon Δ (such as 1 day or 10 days), the loss of the portfolio over time period $[s, s+\Delta]$ is*

$$L[s, s+\Delta] = -[V(s+\Delta) - V(s)].$$

The distribution of $L[s, s+\Delta]$ is the portfolio's loss distribution, we write L_{t+1} as loss from day t to day t+1. (Blais, 2018)

Def 5. *The risk factor changes $(X_t)_{t\in\mathbb{N}}$ by*

$$\underline{X}_{t+1} = \underline{Z}_{t+1} - \underline{Z}_t,$$

then the loss is

$$L_{t+1} = -[V_{t+1} - V_t] = -[f(t+1, \underline{z}_{t+1}) - f(t, \underline{z}_t)].$$

Def 6. *The linearized loss is defined by*

$$L^{\Delta}_{t+1} = -[f_t(t, \underline{z}_t)\Delta t + \sum_{i=1}^{d} f_{z_i}(t, \underline{z}_t) X_{t+1,i}],$$

which is the first order taylor approx, at the point $(t, \underline{z}_t)$ *when evaluated at* $(t+1, \underline{z}_{t+1})$.

3.2.1 Example of Stock Portfolio

Suppose we have d stocks, λ_i is the number of shares in stock i, the risk factor is (Blais, 2018)

$$\underline{Z}_t = [ln(S_{t,1}), ln(S_{t,2}), \cdots, ln(S_{t,d})]^T,$$

where $S_{t,i}$ is the price of stock i at time t. Then

$$L^{\Delta}_{t+1} = -\sum_{i=1}^{d} \lambda_i S_{t,i} X_{t+1,i} \tag{3.1}$$

$$\begin{aligned} L^{\Delta}_{t+1} &= -\sum_{i=1}^{d} \lambda_i S_{t,i} X_{t+1,i} \\ &= -V_t \underline{\omega}_t^T \underline{X}_{t+1}, \end{aligned} \tag{3.2}$$

where weight $\omega_{t,i}$ is defined as

$$\omega_{t,i} = \frac{\lambda_i S_{t,i}}{V_t}.$$

3.2.2 Example of Bond Portfolio

Suppose we have zero coupon bonds at maturity T and we denote the price of bond at time $0 \leq t \leq T$ as $P(t, T)$ which is the present value of that $1 paid at time T.
Def 7. *The continuously compounded yield for this zero coupon bond can be wirtten as: (Blais, 2018)*

$$y(t, T) = -\frac{1}{T-t} ln[P(t, T)]$$

From the definition, we can write the bond price as:

$$P(t, T) = e^{-(T-t)y(t,T)}$$

So then suppose we have a portfolio with d-default free zero coupon bond with maturities T_i & prices $p(t, T_i), i = 1, \cdots, d$ and we denote λ_i to be the number of the bond i. Then the risk factors can be written as:

$$Z_{t,i} = y(\tau_i, T_i), i = 1, \cdots, d$$

here the discrete time $\tau_t = t\Delta t$, the value of the portfolio can be written as:

$$V_t = g(\tau_t, Z_t) = \sum_{i=1}^{d} \lambda_i e^{Z_{t,i}(T_i - \tau_t)}$$

Then we can have:

$$L_{t+1} = -\sum_{i=1}^{d} \lambda_i [P(\tau_{t+1}, T_i) - P(\tau_t, T_i)] \tag{3.3}$$

3.2.3 Example of Option Portfolio

Here we take an European Call option as an example. Suppose our underlying asset is non-dividend-paying stock with price at time s is S_s, the maturity date is T and strike price is K. We have the price fomula from Black-Scholes formula (Turner, 2010) can be written as:

$$C^{BS}(t, S : r, \sigma, K, T) = SN(d_1) - Ke^{-r(T-s)}N(d_2) \tag{3.4}$$

Here:

$$d_1 = \frac{ln(\frac{S}{K}) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$
$$d_2 = d_1 - \sigma\sqrt{T-t}$$

The payoff can be written as: $Payoff(T) = (S_T - k)^+$. We set N(x) to be the standard normal CDF which is: $N(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$. The time units in years, set $\Delta = \frac{1}{250}$.

The risk factor changes can be represented as:

$$X_{t+1} = \begin{bmatrix} ln(S_{t+1}) - ln(S_t) \\ r_{t+1} - r_t \\ \sigma_{t+1} - \sigma_t \end{bmatrix}$$

Then for the loss we have:

$$L_{t+1} = -[C^{BS}([t+1]\Delta, Z_t + X_{t+1}) - C^{BS}(t\Delta, Z_t)] \tag{3.5}$$

For the linearized loss:

$$L_{t+1}^{\Delta} = -[C_t^{BS}\Delta + C_S^{BS} S_t X_{t+1,1} + C_r^{BS} X_{t+1,2} + C_\sigma^{BS} X_{t+1,3}] \tag{3.6}$$

which leads us to 4 Greeks of the option: C_t^{BS} called Theta Θ which measures the sensitivity of the value of the derivative to the passage of time:$\Theta = -\frac{\partial V}{\partial t}$;C_S^{BS} called Delta Δ which measures the rate of change of the theoretical option value with respect to changes in the underlying assset's price:$\Delta = \frac{\partial V}{\partial S}$; C_r^{BS} called Rho ρ which measures sensitivity to the interest rate:$\rho = \frac{\partial V}{\partial r}$; C_σ^{BS} called Vega υ which measures sensitivity to volatility: $\upsilon = \frac{\partial V}{\partial \sigma}$.

Furthermore, we can write our second order loss as following: (Blais, 2018)

$$L_{[t]}^{\Delta\Gamma}(X) = -[C_t^{BS}\Delta + C_S^{BS} S_t X_1 + C_r^{BS} X_2 + C_\sigma^{BS} X_3 + \frac{1}{2}(C_{ss}^{BS} S_t^2 X_1^2 + 2C_{s\sigma}^{BS} S_t X_1 X_3 + C_{\sigma\sigma}^{BS} X_3^2)] \tag{3.7}$$

which leads us to second-order Greeks: C_{ss}^{BS} called Gamma Γ which measures the rate of change in the delta with respect to changes in the underlying price: $\Gamma = \frac{\partial \Delta}{\partial S} =$

$\frac{\partial^2 V}{\partial S^2}$; $C^{BS}_{s\sigma}$ called Vanna which is a second order derivative of the option value, once to the underlying spot price and once to volatility: $Vanna = \frac{\partial \Delta}{\partial \sigma} = \frac{\partial^2 V}{\partial S \partial \sigma}$; $C^{BS}_{\sigma\sigma}$ called Vomma measures second order sensitivity to volatility: $Vomma = \frac{\partial^2 V}{\partial \sigma^2}$.

3.3 Order Statistics and the Sample CDF

3.3.1 Sample CDF and Order statistics

Suppose that $X_1, \cdots, X_n$ is a random sample from a probability distribution with CDF F. In this section we estimate F and its quantiles. The sample or empirical CDF $F_n(x)$ is defined to be the proportion of the sample that is less than or equal to x. More generally,

$$F_n(x) = \frac{\sum_{i=1}^{n} I\{X_i \leq x\}}{n},$$

where I is indicator function.

The order statistics $X_{(1)}, \cdots, X_{(n)}$ are$X_1, \cdots, X_n$ ordered from smallest to largest. The subscripts of the order statistics are in parentheses to distinguish them from the unordered sample. The sample q-sample quantile (100qth sample percentile) is $X_{(k)}$ where k is qn rounded to an integer. (Blais, 2017)

3.3.2 QQ plot

Many statistical models assume that a random sample comes from a normal distribution. Normal probability plots are used to check this assumption. If the assumption is true, then the qth sample quantile will be approximately equal to $\mu + \sigma\Phi^{-1}(q)$, which is the population quantile. Therefore, except for sampling variation, a plot of the sample quantiles versus Φ^{-1} will be linear. The normal probability plot is a plot of $X_{(i)}$ versus $\Phi^{-1}\{\frac{i}{1+n}\}$ (these are the i/(n + 1) sample and population quantiles, respectively). Systematic deviation of the plot from a straight line is evidence of nonnormality. (Blais, 2017)

3.3.3 Heavy-Tailed Distribution

Distributions with high tail probabilities compared to a normal distribution with the same mean and standard deviation are called heavy-tailed. Because kurtosis is particularly sensitive to tail-weight, high kurtosis is nearly synonymous with having heavy tails. Heavy-tailed distribution are important models in finance, because stock return distribution have often been observed to have heavy tails. A heavy-tailed distribution is more prone to extreme values, which are sometimes called outliers. (Blais, 2017)

3.3.4 t-distribution (Student-t)

Density:

$$g_\nu(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu}\,\Gamma(\frac{\nu}{2})[1 + \frac{x^2}{\nu}]^{\frac{\nu+1}{2}}}$$

where $\Gamma(z) = \int_0^\infty e^{-t}t^{z-1}dt$ for $n \in \mathbb{N}$, and $\Gamma(n) = (n-1)!$. (Blais, 2018)

3.4 Risk Measure

3.4.1 Value at Risk

Def 8. *Given condifence level $\alpha \in (0,1)$, the value-at-risk of our portfolio at level α is the smallest number l such that the probability that the loss L exceeds l is no longer than $1-\alpha$, formally*

$$VaR_\alpha = \inf\{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\}.$$

Theorem 1. *For $F_l \sim N(\mu, \sigma^2)$ and $\alpha \in (0,1)$,*

$$VaR_\alpha = \mu + \sigma \mathbb{N}^{-1}(\alpha),$$

where $\mathbb{N}$ is the CDF for the distribution of $N(0,1)$. (Blais, 2018)

3.4.2 Expected Shortfall

Def 9. *For a loss L with $E(|L|) < \infty$ and CDF F_L, the expected shortfall at confidence level $\alpha \in (0,1)$ is*

$$ES_\alpha = CVaR_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u du.$$

Moreover,

- If $F_L \sim N(\mu, \sigma^2)$, then
$$ES_\alpha = \mu + \sigma\frac{\phi[N^{-1}(\alpha)]}{1-\alpha},$$
where ϕ is the PDF and N is the CDF for the N(0,1) distribution.
- Suppose L has a t-distribution, $\tilde{L} = \frac{L-\mu}{\sigma}$ has the standard t distribution with $\nu > 1$. Then if $\nu > 2$, $E[L] = \mu, Var(L) = \frac{\nu}{\nu-1}\sigma^2$, and
$$ES_\alpha = \mu + \sigma ES_\alpha(\tilde{L}),$$
where
$$ES_\alpha(\tilde{L}) = \frac{g_\nu[t_\nu^{-1}(\alpha)]}{1-\alpha}\left(\frac{\nu + [t_\nu^{-1}(\alpha)]^2}{\nu - 1}\right),$$
t_ν is the CDF and g_ν is the PDF of the standard t distribution. (Blais, 2018)

3.5 How far away the actual monthly losses from the VaR and CVaR

Suppose our actual monthly loss is L_t, and value at risk at level α is VaR_α, the corresponding expected shortfall is ES_α, then we focus on

$$StaV_\alpha = \frac{VaR_\alpha - L_t}{VaR_\alpha} \tag{3.8}$$

and

$$StaCV_\alpha = \frac{ES_\alpha - L_t}{ES_\alpha} \tag{3.9}$$

to describe how far away the actual monthly losses from the VaR and CVaR.

In general, VaR_α and ES_α (when α is 0.95 or 0.99) is positive number, L_t could be positive or negative number. When $L_t < 0$, means we get profit from the portfolio, when $L_t > 0$, means we lose money from the portfolio. For equation (3.8) or equation (3.9),

- when $StaV_\alpha \geq 1$ or $StaCV_\alpha \geq 1$, it means $L_t \leq 0$, we earn money, the actual monthly loss is on the left of distribution plot.
- when $0 \leq StaV_\alpha < 1$ or $0 \leq StaCV_\alpha < 1$, it means we lose money, the actual monthly loss is on the right of distribution plot, and still under the range of the distribution.
- when $StaV_\alpha < 0$ or $StaCV_\alpha < 0$, it means we lose money, but the actual monthly loss is beyond the range of the distribution, the actual monthly loss is a extremely value worse than VaR_α or $CVaR_\alpha$.

3.6 Estimation of VaR and CVaR using Polynomial Tails

3.6.1 Methods of Estimation

a. **Parametric method**: works for large α which usually close to 1. But it's sensitive to model mis-specification.

b. **Non-parametric method**: works for small α and requires a large sample size.

c. **Semi-parametric method**: it combines a & b. First use a non-parametric estimate of VaR_{α_0} for small α. Then use a parametric model to estimate VaR_α & $CVaR_\alpha$ for $\alpha > \alpha_0$. And this is the method we use in this project. (Blais, 2018)

Suppose the loss density has a polynomial right tail $f(y) = Ay^{-(a+1)}$ for all $y \geq c$, for some $c > 0$, and $A, a > 0$. Thus we have

$$\begin{aligned} P(L \geq y) &= \int_y^\infty f(u)du \\ &= \int_y^\infty Au^{-(a+1)}du \\ &= \frac{A}{a}y^{-a} \end{aligned} \tag{3.10}$$

Then we suppose $y_0, y_1 > 1$, we have

$$\begin{aligned} \frac{P(L > y_1)}{P(L > y_0)} &= (\frac{y_1}{y_0})^{-a} \\ &= \frac{VaR_{\alpha_1}}{VaR_{\alpha_0}} \\ &= (\frac{1-\alpha_0}{1-\alpha_1})^{\frac{1}{a}} \end{aligned}$$

So we can estimate VaR_α by

$$VaR_\alpha \approx VaR_{\alpha_0}(\frac{1-\alpha_0}{1-\alpha})^{\frac{1}{a}} \tag{3.11}$$

And then ES_α by

$$ES_\alpha \approx \frac{a}{a-1}VaR_\alpha \tag{3.12}$$

3.6.2 Estimate Tail Index

We have two approaches:

1. **Regression Estimator**:

   Recall: For $l > 0$, $P(L \geq l) = \frac{A}{a}l^{-(a)}$. And From this we can get

   $$ln[P(L \geq l)] = ln(\frac{A}{a}) - alnl \tag{3.13}$$

   Then given sample of loses $L_1, L_2, \cdots, L_n$ with the orer stats: $L_{(1)} \leq L_{(2)} \leq \cdots \leq L_{(n)}$.

   We set $L_{(k)}$ which means there are $n-k$ losses that are larger than $L_{(k)}$. We can estimate that $P(L \geq L_{(k)}) \approx \frac{n-k}{n}$ then we get $ln[P(L \geq L_{(k)})] \approx ln(\frac{n-k}{n})$,

   Combined with (3.13) and setting $l = L_{(k)}$, we get

   $$ln(\frac{n-k}{n}) \approx ln(\frac{A}{a}) - aln(L_{(k)})$$

   Rewrite in the order:

   $$ln(L_{(k)}) \approx \frac{1}{a}ln(\frac{A}{a}) - \frac{1}{a}ln(\frac{n-k}{n}) \tag{3.14}$$

   From (3.14) we can plot the points $\{ln(\frac{n-k}{n}), ln(L_{(k)})\}_{k=n-m}^{n}$ for m being a small percentage of n. (Blais, 2018)

   We aim to fit this scatter plot with a line regression and the slope of the regression is an estimator of $-\frac{1}{a}$.

   For instance, in our project we have the following plot for Regression estimator:

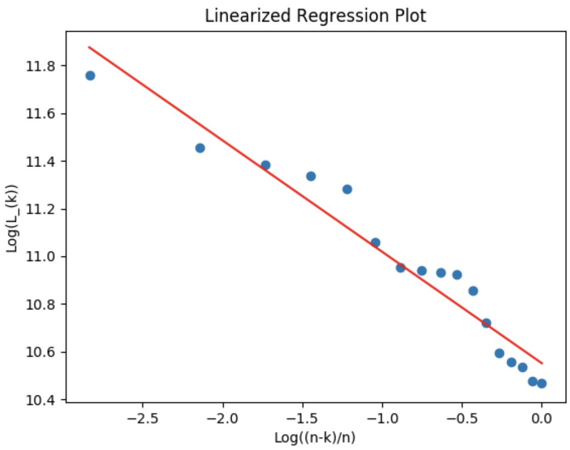


FIGURE 3.1: Example of how we estimate parameter a through Regression method.

From the above plot, we plot the points we need and fit them into a regression line, we can easily get the slope which is $-\frac{1}{a}$, then we can get the estimation of parameter a.

2. **Hill Estimator**:

We let $L_{n,n} \leq L_{n-1,n} \leq \cdots \leq L_{2,n} \leq L_{1,n}$ be the upper order statistics of losses $L_1, L_2, \cdots, L_n$, and $n(c)$ be the number of losses L_i which are larger than c.

Recall the conditional density of L_i given that $L_i \geq c$ is $\frac{ac^a}{l^{a+1}}$. This can lead us to the Likelihood function for $L_{1,n}, L_{2,n}, \cdots, L_{n(c),n}$ is

$$
\begin{aligned}
L_{(a)} &= \prod_{t=1}^{n(c)} \frac{ac^a}{L_{t,n}^{a+1}} \\
&= f_{L_{1,n},\cdots,L_{n(c),n}}(L_{1,n}, \cdots, L_{n(c),n}) \\
&= \prod_{t=1}^{n(c)} f_{L_{t,n}}(L_{t,n})
\end{aligned} \tag{3.15}
$$

We need to choose a that maxmize (3.15) which can be done by maxmize $ln(L_{(a)})$. (Blais, 2018)

$$
\begin{aligned}
ln(L_{(a)}) &= \sum_{t=1}^{n(c)} ln(\frac{ac^a}{L_{t,n}^{a+1}}) \\
&= \sum_{t=1}^{n(c)} [ln(ac^a) - ln(L_{t,n}^{a+1})] \\
&= \sum_{t=1}^{n(c)} [ln(a) + aln(c) - (a+1)ln(L_{t,n})]
\end{aligned}
$$

Then derive it, we get

$$\frac{dln(L_{(a)})}{da} = \sum_{t=1}^{n(c)}[\frac{1}{a} + ln(c) - ln(L_{t,n})]$$
$$= n(c)\frac{1}{a} + \sum_{t=1}^{n(c)} ln(\frac{c}{L_{t,n}})$$

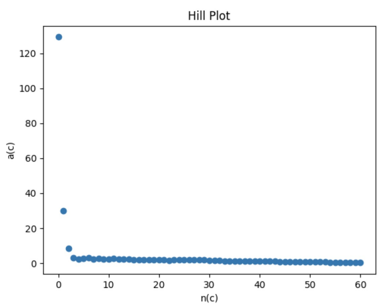
Then we set it to 0, it would give us $n(c)\frac{1}{a} = -\sum_{t=1}^{n(c)} ln(\frac{c}{L_{t,n}})$. So we can get

$$a = \frac{n(c)}{\sum_{t=1}^{n(c)} ln(\frac{L_{t,n}}{c})} \tag{3.16}$$

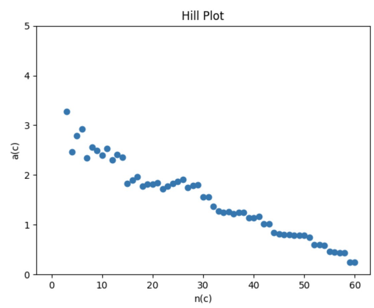
We need to properly choose $n(c)$: If $n(c)$ is too large then $f(l) = Ay^{-(a+1)}$ would likely be a poor fit and lead to model error; If $n(c)$ is too small, there will be too few data points that exceeding c, thus would lead to that our HIll estimator $\hat{a}(c)$ be highly variable.

We should plot $\hat{a}(c)$ versus $n(c)$ and look for a range of values for $n(c)$ where $\hat{a}$ is relatively constant. Thus we choose our $n(c)$ in that range.

In our project part 2, we use weekly data for estimation which is different from part 1. We consider 150 weeks which is almost 3 years would be a good length for back-testing because it can provide us enough data for the estimation and otherwise would create too many useless graphs. In this part of project, we set the estimation of a to be the proper point from the group of points which have the most points that satisfy the condition that the value between them are less than 0.1. In the plot, it come from the most smooth group of points that around to be the same value.



(A) Hillplot with all the points.



(B) enlarged Hillplot in order to see the trend.

FIGURE 3.2: Hillplot using data from 2013/12/31 to 2016/12/31.

For example, the above figure shows the Hillplot using data from 2013/12/31 to 2016/12/31. From part B we would get our estimation of a would be around 1.8.

3.6.3 Historical Simulation

The historical Simulation is very popular among banks for trading book[1] which means the portfolio of financial instruments held by a brokerage or bank. The idea is that we use the historical data $X_{t-n+1}, \cdots, X_t$ which has an empirical distribution to estimate F_L.

From the time series $\{\tilde{L_s} = l_{[t]}(\underline{X}_s) : s = t-n+1, \cdots, t\}$, we build the empirical CDF: $F_n(l) = P(L \leq l) = \frac{1}{n}\sum_{s=t-n+1}^{t} \mathbb{1}_{\tilde{L_s} \leq l}$. If the data $\underline{X}_s$ are identical independent distribution then from $F_n(l)$ we can get $F_L(l)$. (Blais, 2018)

3.7 GARCH Model

3.7.1 ARMA(p,q) Time Series

Autoregressive moving average (ARMA) models are useful statistical tools to examine the dynamical characteristics of ecological time-series data. (Lves AR, 2010) We could use it to model for how y_t deviates from its mean based on prior information and a new noise term ε_t as following:

$$y_t - \mu = \varphi_1(y_{t-1} - \mu) + \cdots + \varphi_p(y_{t-p} - \mu) - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q} + \varepsilon_t \tag{3.17}$$

Def 10. *Let X be a square integrable real-valued random variable defined on probability space* $(\Omega, \mathcal{F}, P)$ *and g be a solo-*σ*-algebra of* F*. Then the conditional variance of X given g is written as* $Var(X \mid g) = E[(x - E[x \mid g])^2 \mid g]$*. The condition variance of X given random variable y defined on* $(\Omega, \mathcal{F}, P)$ *is written as* $VaR(X \mid y) = E[(x - E[x \mid y])^2 \mid y]$*.*

3.7.2 ARCH(1) Process

Def 11. *A finite time series exhibits heteroscedasticity if it has non-constant variance. A time series exhibits conditional heteroscedasticity if it has a non-constant conditional variance.(White, 1980)*

Def 12. *The process* a_t *is said to be an ARCH(1)* [2] *process if* $a_t = \varepsilon_t\sqrt{\alpha_0 + \alpha_1 a_{t-1}^2}$ *where* $\alpha_0, \alpha_1 \geq 0$ *&* $\alpha_1 < 1$ *so that* $Var(a_t) < \infty$ *&* a_t *is stationary. (Nelson, 1991)*

Through observation, we notice that if we square both side of the defining equation we get: $a_t^2 = \varepsilon_t^2(\alpha_0 + \alpha_1 a_{t-1}^2)$ which like an AR(1) model for a_t^2 with multiplicative rather than additive noise ε_t.

We can get the conditional expectation as:

$$\begin{aligned}
E[a_t \mid a_{t-1}, a_{t-2}, \cdots] &= E[\varepsilon_t\sqrt{\alpha_0 + \alpha_1 a_{t-1}^2} \mid a_{t-1}, a_{t-2}, \cdots] \\
&= \sqrt{\alpha_0 + \alpha_1 a_{t-1}^2} E[\varepsilon_t \mid a_{t-1}, a_{t-2}, \cdots] \\
&= \sqrt{\alpha_0 + \alpha_1 a_{t-1}^2} E[\varepsilon_t] \\
&= 0
\end{aligned} \tag{3.18}$$

[1] `https://www.investopedia.com/terms/t/tradingbook.asp`

[2] ARCH in this article stands for Auto-Regressive Conditional Heteroscedasticity.

We can also get the conditional variance as:

$$\begin{aligned}\sigma_t^2 &= Var(a_t \mid a_{t-1}, a_{t-2}, \cdots) \\ &= E[(a_t - E[a_t \mid a_{t-1}, a_{t-2}, \cdots])^2 \mid a_{t-1}, a_{t-2}, \cdots] \\ &= E[a_t^2 \mid a_{t-1}, a_{t-2}, \cdots] \\ &= (\alpha_0 + \alpha_1 a_{t-1}^2) E[\varepsilon_t^2 \mid a_{t-1}, a_{t-2}, \cdots] \\ &= (\alpha_0 + \alpha_1 a_{t-1}^2) E[\varepsilon_t^2] \\ &= \alpha_0 + \alpha_1 a_{t-1}^2 \end{aligned} \tag{3.19}$$

From above we can see if a_{t-1} is large, then a_{t-1}^2 is large, then σ_t^2 is large, we can say a_t is expected to have a large deviation from its mean which is zero.

Take the expectation of both side of (3.19), we have:

$$\begin{aligned}E(\sigma_t^2) &= E[\alpha_0 + \alpha_1 a_{t-1}^2] \\ &= E[Var(a_t \mid a_{t-1}, \cdots)] \\ &= \alpha_0 + \alpha_1 E[a_{t-1}^2]\end{aligned}$$

Then use the g-σ-algebra rule which is $E[E(X \mid g)] = E[X]$, we gan derive from the above equation as:

$$\begin{aligned}\sigma &= \alpha_0 + \alpha_1 \sigma^2 \\ &= \frac{\alpha_0}{1 - \alpha_1}\end{aligned} \tag{3.20}$$

Notice that here σ^2 is the unconditional variance of a_t which is different from σ_t^2.

We denote $\gamma(h)$ as auto covariance function; $\rho(h)$ as auto correlation function; $\gamma_a(h) = Cov(a_t, a_{t-h})$; $\rho_a(h) = Corr(a_t, a_{t-h})$. Then for the process a_t, we have: $\gamma_a(0) = Var(a_t) = \sigma^2$; $\gamma_a(h) = 0$ for $h \neq 0$; $\rho_a(h) = 0$ for $\forall h \neq 0$, $\rho(h) = \varphi^{|h|}$; $\rho_{a^2}(h) = \alpha_1^{|h|}$ where $0 \leq \alpha_1 < 1$ and $|\varphi| < 1$. (R. Rabemananjara, 1993)

3.7.3 GARCH(p,q) Process

Def 13. *Let $\varepsilon_t \sim WN(0,1)$, a_t is said to be an ARCH(q) process if $a_t = \sigma_t \varepsilon_t$ and $\sigma_t = \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i a_{t-i}^2}$. (Blais, 2018)*

Def 14. *Let $Z_t \sim WN(0,1)$, the process X_t is a GARCH(p,q) [3] process if it's stationary and satisfies $\forall t \in Z$ & some strictly positive process σ_t, $X_t = \sigma_t Z_t$. Here $\sigma_t = \sqrt{\alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2}$, where $\alpha_0 > 0$, $\alpha_i \geq 0$ for $i = 1, \cdots, p$ & $\beta_j \geq c$ for $j = 1, \cdots, q$. (Engle, 1996)*

Theorem 2. *The GARCH(1,1) process is a weak white noise and we can say $\alpha_1 + \beta_1 < 1$, it's variance is $\sigma^2 = \frac{\alpha_0}{1-\alpha_1-\beta_1}$ (Blais, 2018)*

In our project work, we mainly focus on the GARCH(1,1) process. Let's assume $X_t \sim GARCH(1,1)$ with $\alpha_1 + \beta_1 < 1$. Then we can get some useful properties:

1. $E(X_t) = 0$

[3] GARCH in this article stands for Generalized Auto-Regressive Conditional Heteroscedasticity.

2. $E(X_t^2) = \sigma^2$
3. $E(X_{t+h}^2) = \alpha_0 \sum_{i=0}^{h-1}(\alpha_1 + \beta_1)^i + (\alpha_1 + \beta_1)^{h-1}(\alpha_1 X_t^2 + \beta\sigma_t^2)$ and as $h \to \infty$, it converges to σ^2 which is the unconditional variance. (Blais, 2018)

3.7.4 ARMA(1,1)-GARCH(1,1) Process

Def 15. *Let Z_t be WN(0,1). X_t is an ARMA(1,1)-GARCH(1,1) Process if it satisfies following conditions:*

1. $X_t - \mu_t = \varepsilon_t$ *where* $\varepsilon_t = \sigma_t Z_t$
2. $\mu_t = \mu + \varphi(X_{t-1} - \mu) + \theta(X_{t-1} - \mu_{t-1})$
3. $\sigma_t^2 = \alpha_0 + \alpha_1(X_{t-1} - \mu_{t-1})^2 + \beta\sigma_{t-1}^2$ *as* $\alpha_0, \alpha_1, \beta \geq 0$ *&* $\alpha_1 < 1, |\varphi| < 1$ *(Blais, 2018)*

we can write as:

$$X_t - \mu = \varphi(X_{t-1} - \mu) + \theta(X_{t-1} - \mu_{t-1}) + \varepsilon_t \tag{3.21}$$

Then we can get some useful properties:

1. $VaR(X_t) = \frac{\alpha_0}{1-\alpha_1-\beta}$
2. $E[X_{t+h} \mid F_t] = \mu + \varphi^h(X_t - \mu) + \varphi^{h-1}\theta\varepsilon_t$
3. $Var(X_{t+h} \mid F_t) = \alpha_0 \sum_{i=0}^{h-1}(\alpha_1 + \beta)^i + (\alpha_1 + \beta)^{h-1}(\alpha_1\varepsilon_t^2 + \beta\sigma_t^2)$ (C. He, 1999)

The idea is that we use this to predict the σ_{t+h}, μ_{t+h} we need. We need to fit the model to data: typically using max likelihood estimation to obtain the parameter estimates: $\hat{\mu}, \hat{\varphi}, \hat{\theta}, \hat{\alpha_0}, \hat{\alpha_1}, \hat{\beta}$.(A. Thavaneswaran, 1988) We will talk more in the next subsection.

3.7.5 Conditional RIsk Management

In this subsection, we take stock as example to show how GARCH is used for conditional risk management.

We let the value of the assets as V_t, risk factoe change which here is the log-return fot a stock as X_{t+1}. We can represent the linearized loss for the next day as: $L_{t+1}^{\Delta} = l_{[t]}^{\Delta}(X_{t+1}) = -V_t X_{t+1}$.

Our goal is to estimate VaR & ES for the conditional loss distribution. We have the information at time t: $F_t = \sigma(\{X_s : s \leq t\})$ and CDF: $F_{L_{t+1}|F_t}$. Assume WLOG that $V_t = 1$, then we have $L_{t+1}^{\Delta} = -X_{t+1}$

Assume L_t^{Δ} is adapted to F_t and follows a stationary model of the form: $L_t = \mu_t + \sigma_t Z_t$ where μ_t, σ_t are F_{t-1}-measurable & $Z_t \sim WN(0, 1)$.

Then assume it satisfies ARMA(1,1)-GARCH(1,1) and let G be the CDF of the innovations Z_t. Then we can we can write:

$$F_{L_{t+1}|F_t}(l) = G(\frac{l - \mu_{t+1}}{\sigma_{t+1}})$$

(Blais, 2018). And the VaR calculation would give:

$$VaR_{\alpha}^{t} = \mu_{t+1} + \sigma_{t+1} q_{\alpha}(Z)$$

$$ES_{\alpha}^{t} = \mu_{t+1} + \sigma_{t+1} ES(Z)$$

where Z has the CDF as G, q_{α} means the α-quantile. Specifically, if $G \sim N(0,1)$ then $G(x) = N(x)$, $q_{\alpha}(Z) = N^{-1}(\alpha)$ and $ES_{\alpha}(Z) = \frac{\varphi[N^{-1}(\alpha)]}{1-\alpha}$ where φ is the PDF of N(0,1).

Generally speaking, we first find the historical data $L_{t-n+1}, \cdots, L_t$, then use them to fit the ARMA-GARCH model using max likelihood estimation method, and form the $\hat{\sigma_{t+1}}, \hat{\mu_{t+1}}$ & calculate the $VaR_{\alpha}^{t}, ES_{\alpha}^{t}$ using the formula above.

3.8 Stress Test

Here the Stress Test basically means stress testing in Value at Risk(VaR). A stress test involves running simulations under crises for which a model was not inherently designed to adjust. (Cyril Coste, 2009) There are different methods for calculating the VaR, such as Monte Carlo Simulations (Commonly used in Bloomberg), historical simulations and parametric VaR, that can be stress tested in different ways.

In a historical scenario, the assets you are holding are run through a simulation based on a previous crisis. Examples of historical crises such as 2008-Lehman Default, 2010-Oil prices drop, 2011-Debt Ceiling Crisis & Downgrade.

Chapter 4

Implementation

At the beginning, we use the data from 1/03/2012 to 12/31/2013 as our training data to calibrate three risk models which are fitted monthly loss data by assuming risk factor log-returns follow normal distribution, assuming whole monthly loss follows t distribution and also fitted monthly linearized loss data by assuming risk factor log-returns follow normal distribution. At the meantime, we use the data from 1/02/2014 to 12/31/2017 as our test data for back-testing. Then we use rolling window to estimate each month loss distribution from 1/02/2014 to 12/31/2017, every time we use 2 years historical data as training data and remaining data as testing data. At the end, we analyze risk by value at risk and expected shortfall respectively. We simply choose equally weight $\omega_{t,i} = 1/15$ when forming our portfolio, and suppose our initial capital is $100,0000.

Our whole daily portfolio value process V_t and the daily log-return is

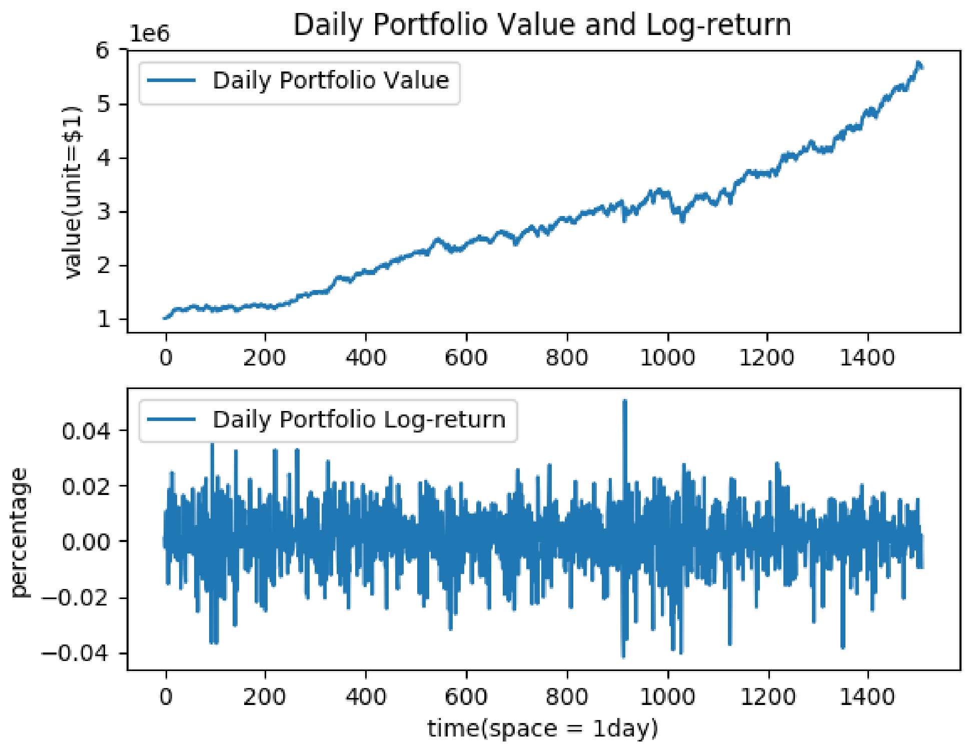
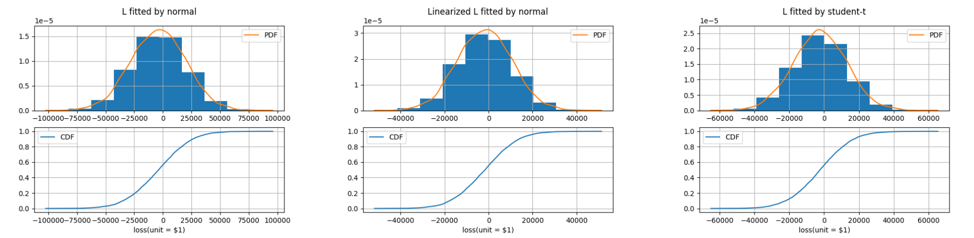


FIGURE 4.1: Whole daily portfolio value process V_t and the daily log-return.

4.1 Calibrate models for estimating the first month's loss distribution of 2014

We first estimate the mean μ and the covariance matrix Σ of the daily log-returns from 1/03/2012 to 12/31/2013, and then we assume that the log-returns of the 15 stocks follow a $N(\mu, \Sigma)$ distribution for next month which is 1/02/2014 - 1/31/2014. By this assumption, we can employ the equation (3.1) and equation (3.2) (for t distribution case, we simply assume the next month's loss distribution follows t distribution)to estimate the next month's loss distribution, and we plot PDF and CDF for the three models respectively as shown in figure 4.2.



(A) Monthly loss fitted by normal. (B) Linearized monthly loss fitted by normal. (C) Monthly loss fitted by t-distribution.

FIGURE 4.2: PDF and CDF of the first month's distribution of 2014.

4.2 Back-test

First we plot the time series for the daily portfolio value process V_t and the log-return for the subsequent years of data from 1/02/2014 to 12/29/2017, it is shown in figure 4.3.

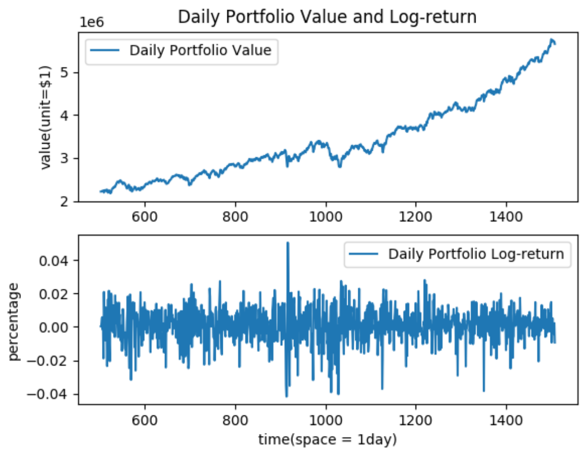


FIGURE 4.3: Daily portfolio value process V_t and the daily log-return from 1/02/2014 to 12/29/2017.

Then we compute the actual losses for each month in the subsequent years, and also plot a curve actual monthly loss versus time t, and a histogram of the actual monthly losses, as shown in figure 4.4.

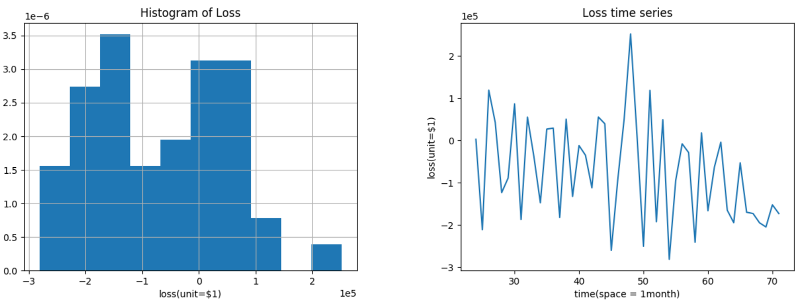


FIGURE 4.4: Histogram and curve of actual monthly loss from 1/02/2014 to 12/29/2017.

By using 2-years rolling window, we can estimate each month's loss distribution, and we compare the actual monthly losses of subsequent years with those corresponding estimated monthly distributions. We add a red vertical line to show the actual monthly loss in both the PDF and CDF distributions for each subsequent months. We only present the 25th(2/28/14), 30th (7/31/2014), 50th (3/31/2016), 60th (1/31/2017) and 70th (11/30/2017) months' distribution in figure 4.6, others are shown in appendix **??**. Note that in figure 4.6, figure (A), (B), and (C), are the 25th distributions, figure (D), (E), and (F), are the 30th distributions, figure (G), (H), and (I), are the 50th distributions, figure (J), (K), and (L), are the 60th distributions, figure (M), (N), and (O), are the 70th distributions.

We can see from figure 4.6 that the monthly losses come from our different loss distribution are reasonable, since in most cases, the actually monthly losses are under the range of distributions, and by QQ-plot which is the figure 4.5, we can conclude that the empirical quantile and theoretical quantile are very close. But in other hand, it also shows that normal distribution is a too ideal model to used to fit the data.

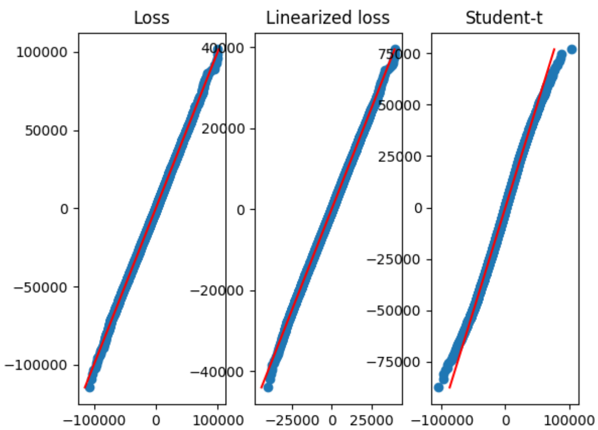
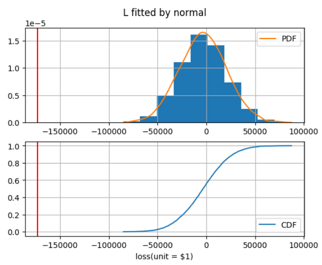
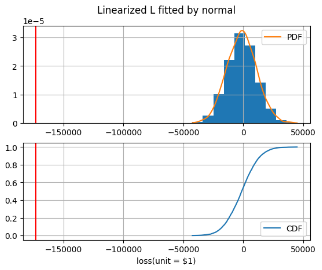


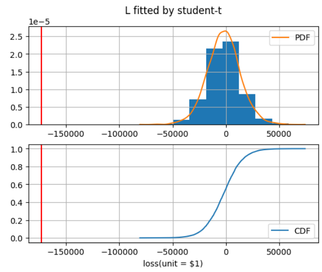
FIGURE 4.5: QQ-plot.



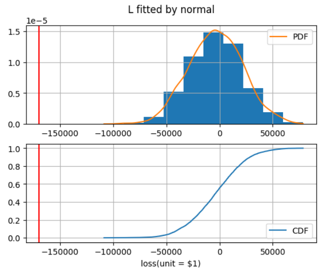
(A) Monthly loss fitted by normal.



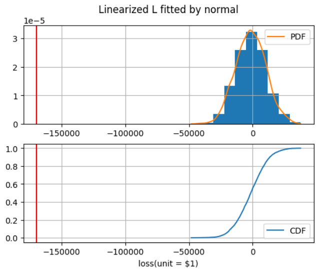
(B) Linearized monthly loss fitted by normal.



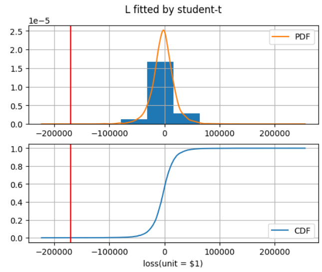
(C) Monthly loss fitted by t-distribution.



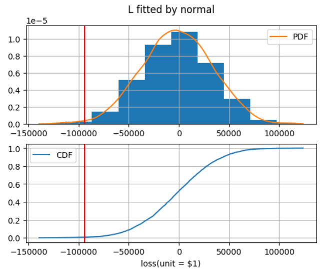
(D) Monthly loss fitted by normal.



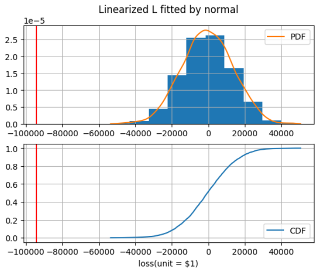
(E) Linearized monthly loss fitted by normal.



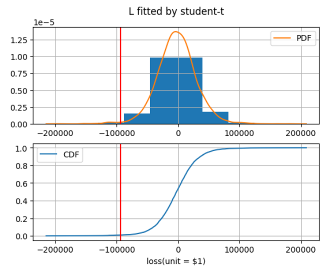
(F) Monthly loss fitted by t-distribution.



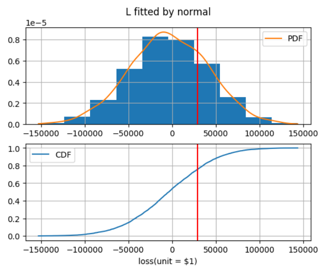
(G) Monthly loss fitted by normal.



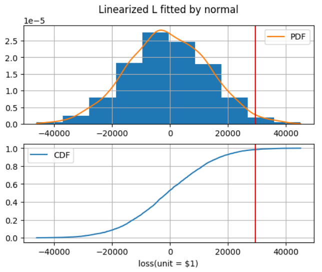
(H) Linearized monthly loss fitted by normal.



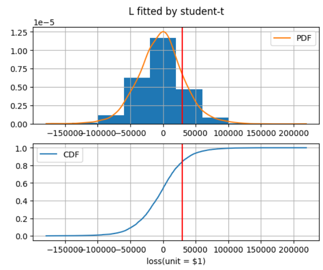
(I) Monthly loss fitted by t-distribution.



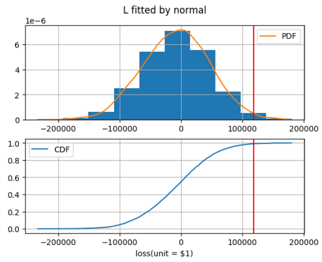
(J) Monthly loss fitted by normal.



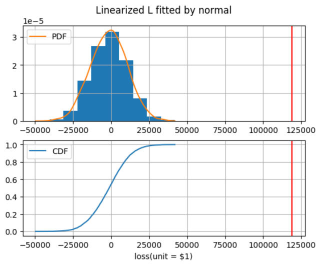
(K) Linearized monthly loss fitted by normal.



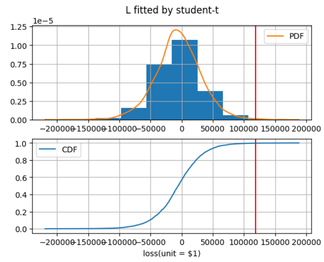
(L) Monthly loss fitted by t-distribution.



(M) Monthly loss fitted by normal.



(N) Linearized monthly loss fitted by normal.



(O) Monthly loss fitted by t-distribution.

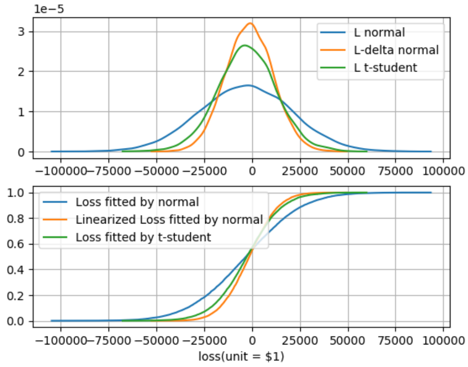
FIGURE 4.6: Estimated loss distribution and realized losses.

4.3 Comparison between normal distribution and t-distribution

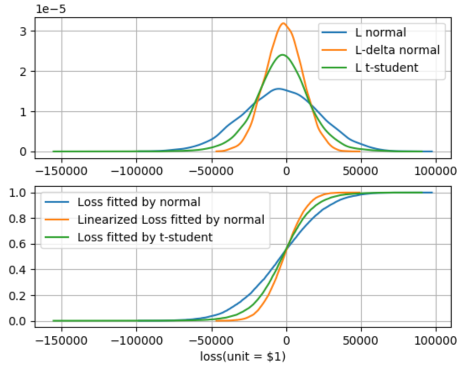
We plot CDF and PDF of three models in a same figure, We only present the 25th (2/28/14), 30th (7/31/2014), 50th (3/31/2016), and 60th (1/31/2017) months' distribution in figure 4.7, others are shown in appendix **??**.

We can see the major difference of them are distribution range (which is the x-axis: loss) and the shape. Overall, the linearized loss distribution fitted by normal(L-delta) is sharpest than others, the t-distribution is moderate, however the loss distribution fitted by normal is the most flatted among the three. On the other hand, the range of distribution of L normal and t-student are much larger than L-delta, and in most situation, t-student's range is larger than L normal which means it consider more extremely situations and has a long tail. The similarity is the mean of the three are same, around zero. And from the figure 4.13, we can see the VaR and CVaR of linearized loss distribution fitted by normal is too small to reflect the reality monthly loss. Most situations the actual monthly loss is beyond the VaR and CVaR at level 0.99. For the other two distributions, they both reflect the actual monthly loss to some extent. The CVaR considers more about the tail, and includes some very extremely cases happened in reality while the VaR could not.

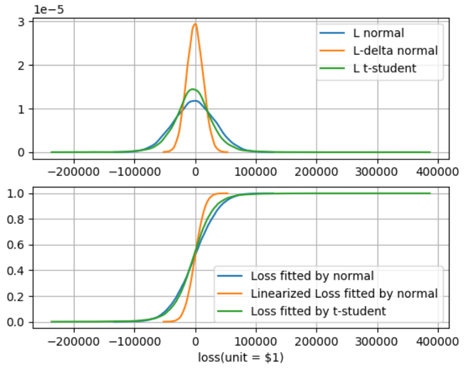
In practical, we would like to use t-distribution to fit our data and combine VaR and CVaR to measure portfolio's risk, since t-distribution is a more flexible method which has 3 parameters to be calibrate while normal distribution only has two. The degree of freedom of t-distribution can have a better fit to our data comparing with normal distribution. And VaR is easy to calculate, we can use VaR firstly help us to save time and money, and use CVaR to get a closer risk measure.



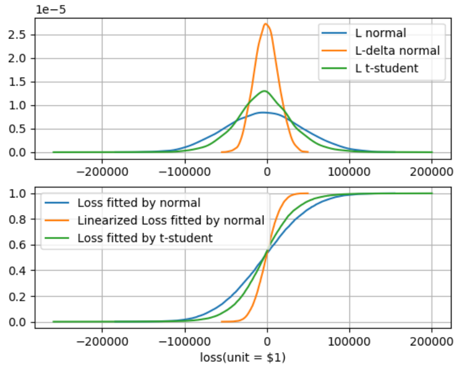
(A) 2/28/2014.



(B) 7/31/2014.



(C) 3/31/2016.



(D) 1/31/2017.

FIGURE 4.7: CDF and PDF of three models.

4.4 Comparison the estimation of a between Hill estimator and regression estimator

In our project part 2, as mentioned in methodology chapter, we use weekly data for our estimation of parameter a. At first, we use the data from 2016/12/31 to 150 weeks before this time which is around 2013/12/31 for our estimation. We get our estimation of a from the two methods. Then we add one week's new data after 2016/12/31 all the way to the week before the last week of 2017. We would get 51 pairs of estimation results by the two different methods. We get our following plot showing the comparison results:

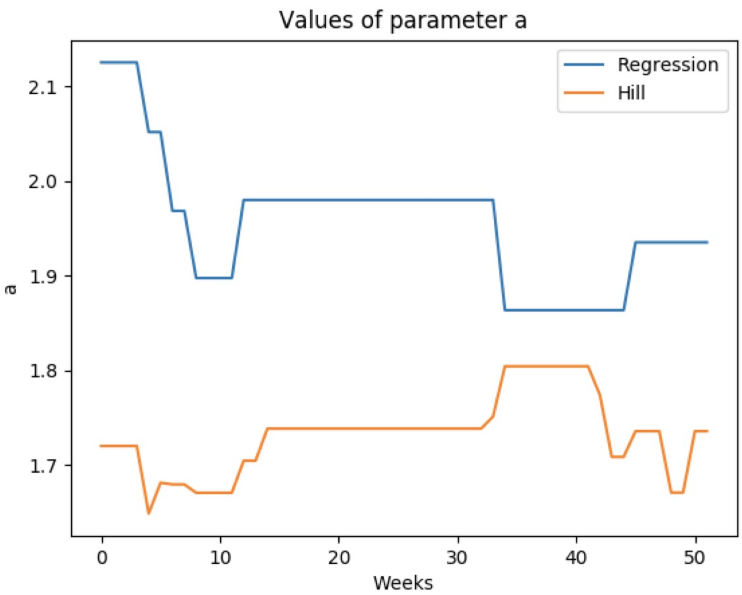


FIGURE 4.8: Comparision of parameter a between Hill estimator and Regression estimator.

From the above plot, we can see the estimation results between the two methods are not the same all the time. And the estimation by Regression method is more stable than Hill estimator because the stable line is longer. In these results there always a gap between the two estimation of a, the estimation of Regression is always greater than the one from Hill estimation. In our project work, the 1st back-testing result between the two method is the largest. From this plot, we somehow believe the Regression method would be a more proper way to estimate the parameter a.

Then we get the following figures showing at confidence level 0.95 and 0.99, how the VaR differs from the two methods:

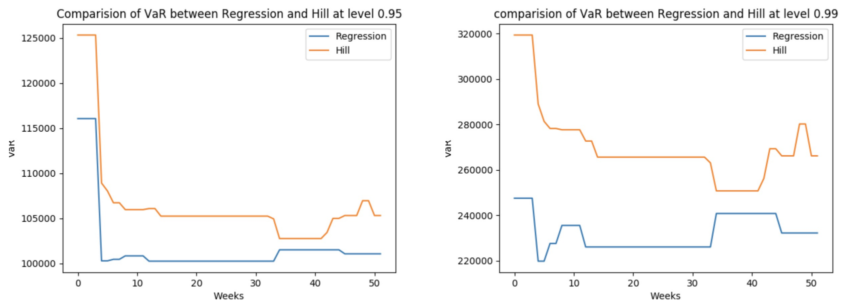


FIGURE 4.9: Comparation of VaR between Regression and Hill method.

From above figures, we can see the Hill estimator method always gives us a higher VaR. The trend of VaR form the two different method is basicly the same. And we still believe Regression estimator is more proper when estimate the parameter a.

Furthermore, we plot the ES as well as CVaR from two methods at confidence level 0.95 and 0.99.

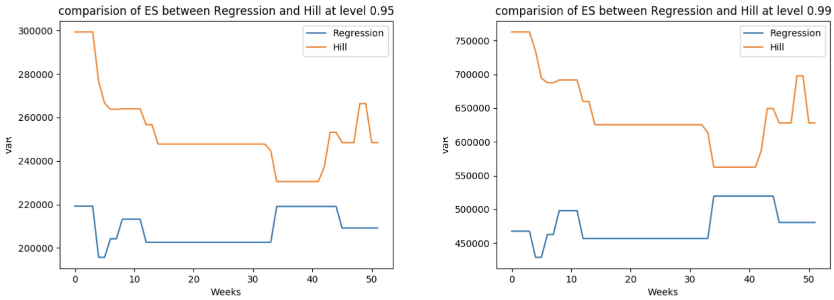


FIGURE 4.10: Comparation of ES between Regression and Hill method.

From above figures, we can get the similar results from the VaR condition which in the same way tell us that in our case Regression estimator is better.

4.5 Using ARMA-GARCH model for Implementation

4.5.1 Comparisons between two versions of GARCH model

We pick weekly data from 2012/01/04 to 2014/12/31 as our historical data to fit the GARCH model for both normal distribution and student t-distribution. Then we get the conditional mean and variance for each week and then plot them for comparision.

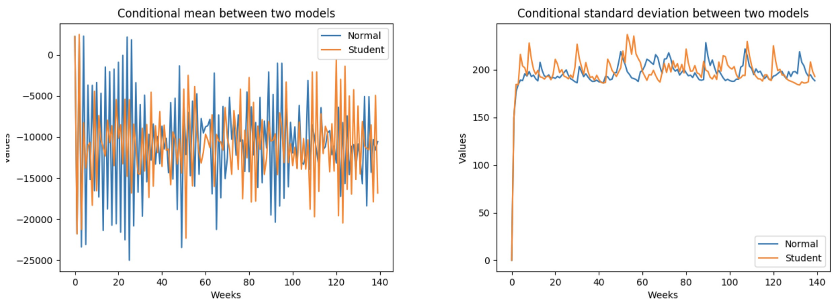
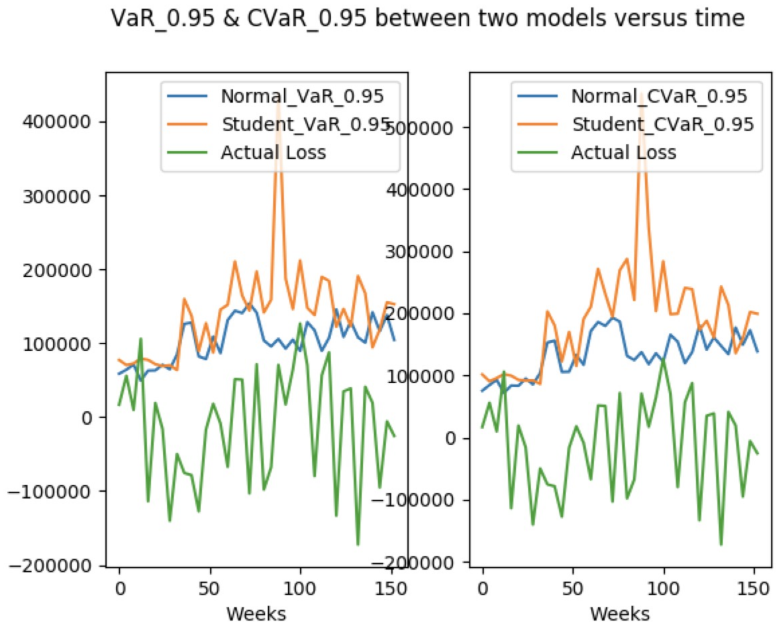


FIGURE 4.11: Comparation of Conditional mean & variance between two models.

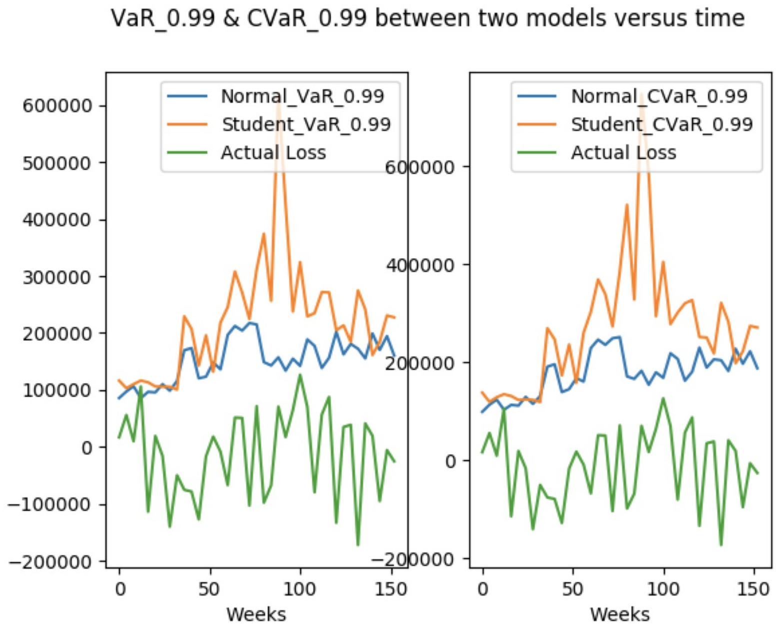
In 4.11, we can see the results of the μ_t & σ_t are on the same level between the two versions of GARCH model. But more precisely, we can see the frequency of the fluctuation given by t-distribution is more than the one given by normal distribution especially for the standard deviation.

4.5.2 VaR and ES for conditional loss distribution

In this part, we plot the figures showing at the confidence level of 0.95 and 0.99 the differnce among the VaR, ES and the actual loss of the portfolio. We are using the weekly data form 2015/01/02 to 2017/12/29 in order to keep track with previous project work. The results are showing from the following figures.



(A) $VaR_{.95}$ and $CVaR_{.95}$.



(B) $VaR_{.99}$ and $CVaR_{.99}$.

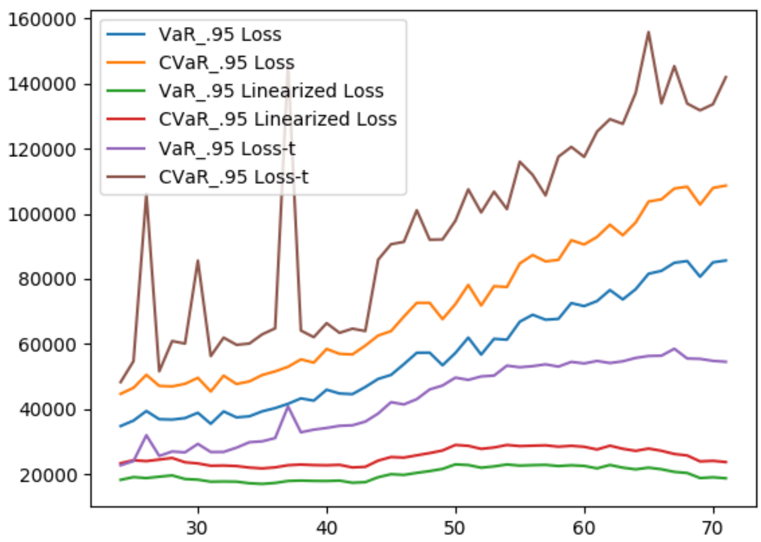
FIGURE 4.12: Comparation of VaR & CVaR between two models

From 4.12, compared to the actual loss, we can see at a 0.95 confidence level sometimes both two methods cannot capture the all the actual loss, but student distribution can always capture more loss than the normal distribution does. At a 0.99 confidence level, still sometimes the results given by the normal distribution is lower than the actual loss, but we can see the student distribution can capture all the losses which indicates student distribution is a better method. But still, we believe both method are good because only one or two times the actual loss exceed the VaR or CVaR we get.

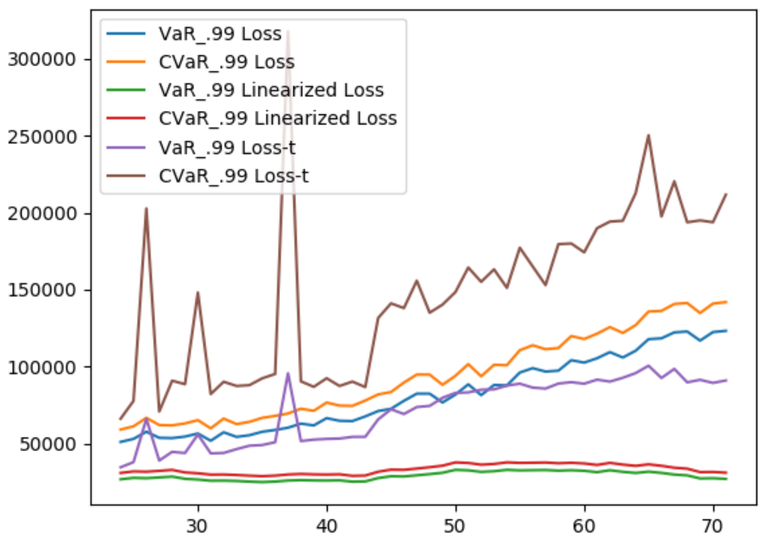
4.6 Risk Reporting

4.6.1 Project part 1

We create a plot of the portfolio VaR_α and $CVaR_\alpha$ versus time for the backtesting range for both the t distribution case and the normal distribution case as shown in figure 4.13.



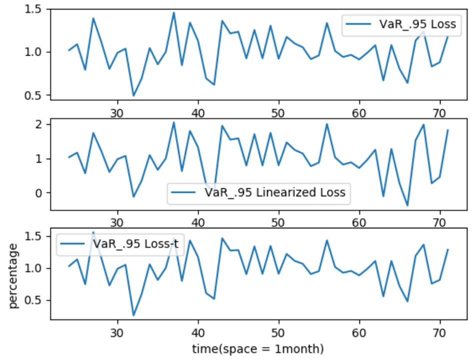
(A) $VaR_{.95}$ and $CVaR_{.95}$.



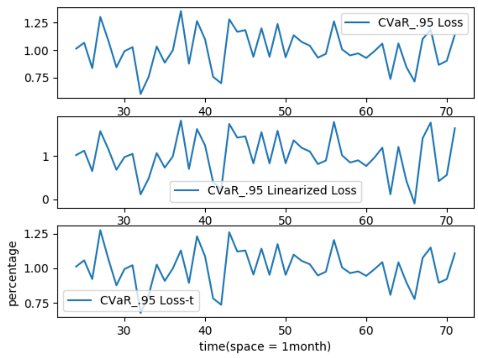
(B) $VaR_{.99}$ and $CVaR_{.99}$.

FIGURE 4.13: VarR and CVaR

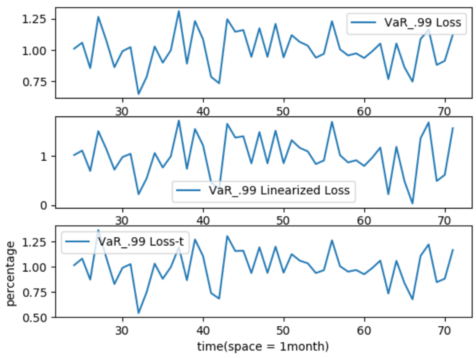
By using the equation (3.8) and equation (3.9), we can plot several pictures to display how far the actual monthly losses from the VaR and $CVaR$ as shown in figure



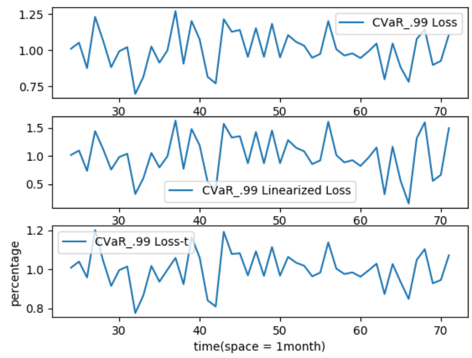
(A) $StaV_{0.95}$.



(B) $StaCV_{0.95}$.



(C) $StaV_{0.99}$.



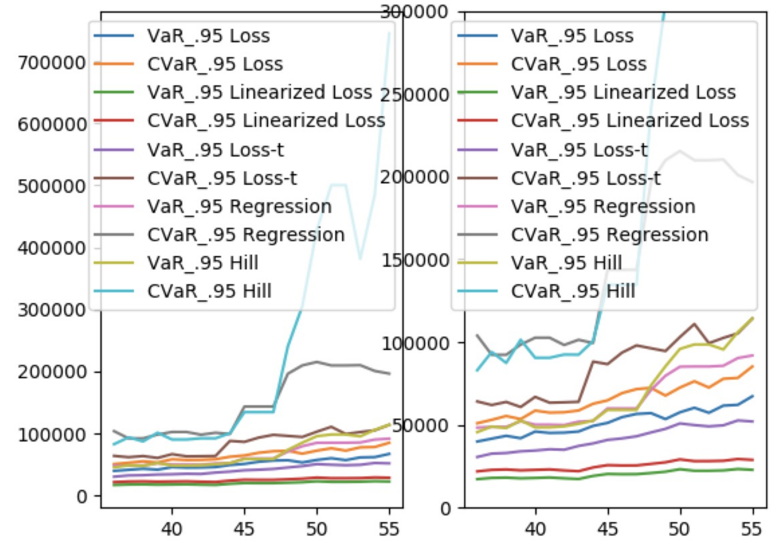
(D) $StaCV_{0.99}$.

FIGURE 4.14: $StaV$ and $StaCV$.

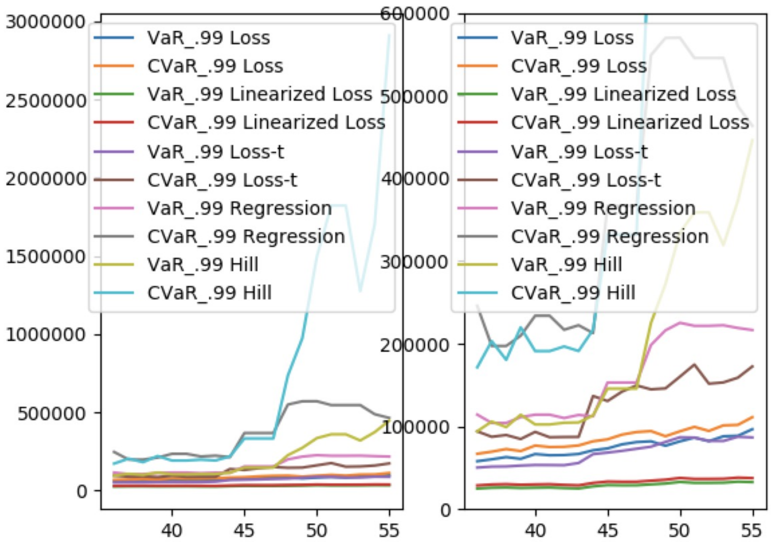
Based on these figures, we can see our StaV and StaCV varies a lot from time to time. The general trends of StaV & StaCV at the same level of α are the same, but StaCV sometimes bigger than StaV when they are both small. Most of the months these values are above 1 which means our portfolio earned money but still we can tell our portfolio had a high volatility. The figures above does not show a long period that those values are always above 1. Overall, our portfolio is somehow risky and simply equal weighted is not an optimal investment strategy.

4.6.2 Project part 2

In this part, compared to part 1, we add the distributions from Hill estimator and Regression estimator methods for further risk reporting. Compared to figure 4.13, we now have the following:



(A) $VaR_{.95}$ and $CVaR_{.95}$.

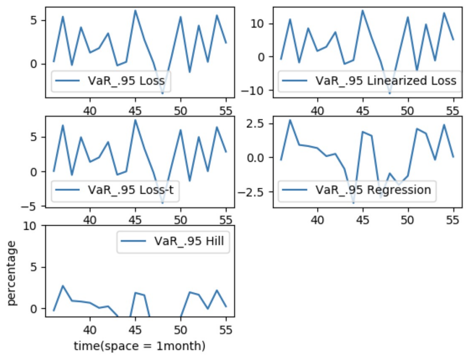


(B) $VaR_{.99}$ and $CVaR_{.99}$.

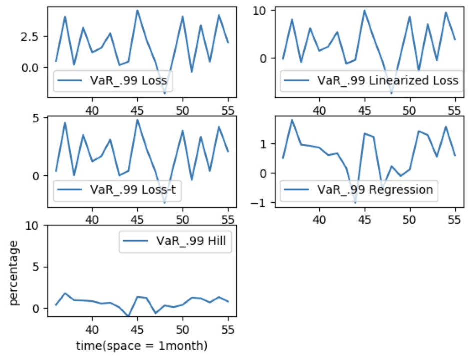
FIGURE 4.15: VaR and CVaR contained Hill and Regression method

In the figure 4.15, the difference between the left and right is that we enlarge the Y-axis in order to observe the difference among lines more clearly. From the figure, we can easily find out that Hill method always gives us bigger VaR and CVaR and sometimes is way too far from other methods. Somehow the Regression method is more consist with other methods.

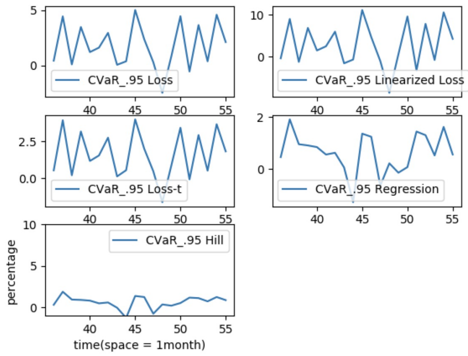
By using the equation (3.8) and equation (3.9), we can plot several pictures to display how far the actual monthly losses from the VaR and $CVaR$ as shown in figure 4.16



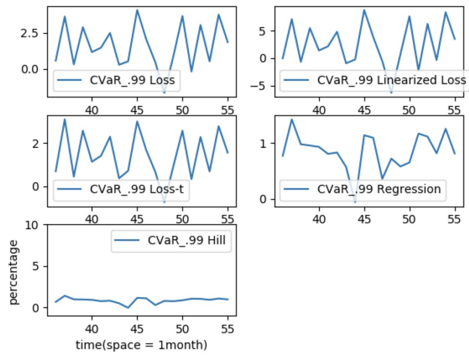
(A) $StaV_{0.95}$.



(B) $StaCV_{0.99}$.



(C) $StaV_{0.95}$.



(D) $StaCV_{0.99}$.

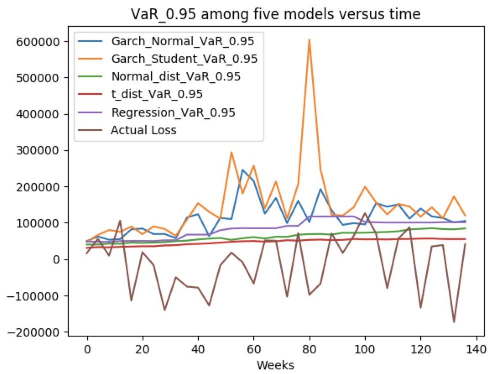
FIGURE 4.16: $StaV$ and $StaCV$.

Compared to figure 4.14 in part 1, when doing this part of risk report, we use the data from 12/31/14 to 1/4/16 for back-testing. Because based on figure 4.4 which shows the monthly loss, we can see during 2017, the monthly losses are all negative which means our portfolio is earning money. There would be no bigger value than the VaR we estimated, so it's useless to capture StaV using this period of data. But during 2014 to 2016 it occured some big losses, so we want to use this period of data to see if our prediction can capture such big losses. If the losses are less than our estimated VaR, than it means our model is proper.

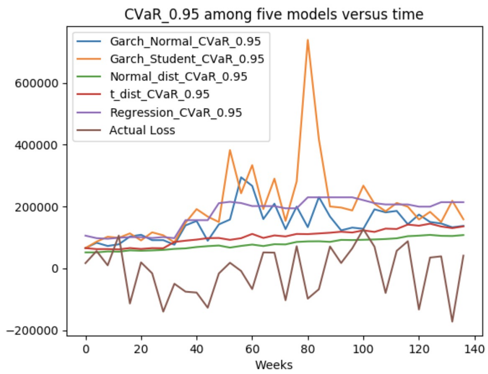
From figure 4.16, we know when the value of Y-axis is below 0, it represents our real loss of the portfolio is bigger than the estimated VaR or CVaR we have in our models. And if the line is smooth that means the difference between the estimated value and the real value is big which indicates a bad model. So in general, HIll model is not a good model compared to others. So when doing parameters estimation, we would better choose to use Regression method.

4.6.3 Project part 3

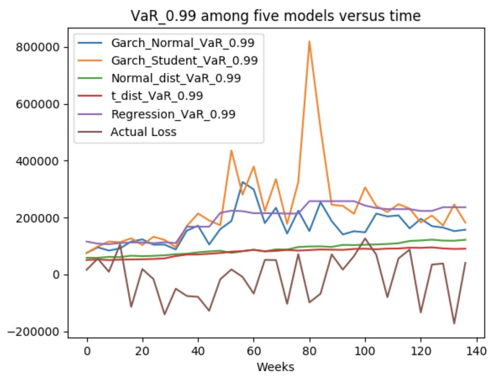
In this part, we add GARCH model into our discussion. Specifically we fit the normal distribution and student distribution with GARCH and compared with other methods for risk reporting. We get the following figures:



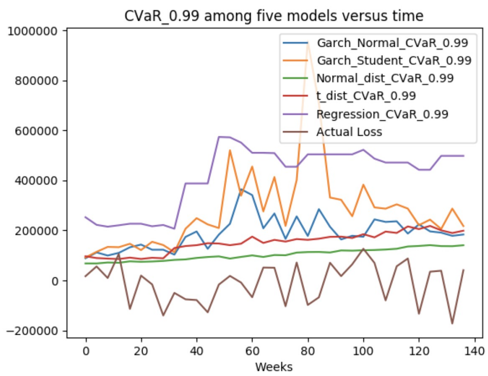
(A) $VaR_{.95}$ among five models.



(B) $CVaR_{.95}$ among five models.



(C) $VaR_{.99}$ among five models.

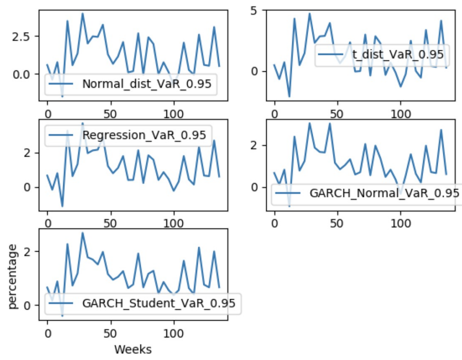


(D) $CVaR_{.99}$ among five models.

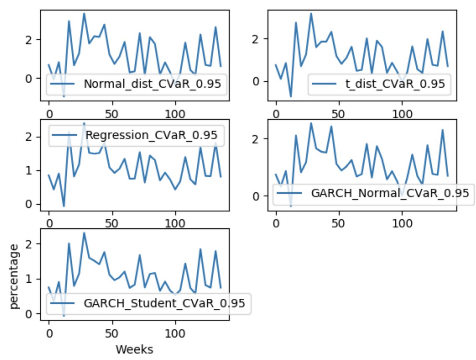
FIGURE 4.17: Comparison of VaR and CVaR among five models.

From 4.17, we add the actual loss for better comparison. we remove the linearized loss method and the hill estimator because we believe the results are not good given by those two and it's better to remove them in order to see the results clear. We can notice that GARCH model and regression method are better than the others because they better capture the bigger loss than the others. The results from other method are too smooth for the fact that the actual loss really varies much from each week. And among the three good method we find out that GARCH based on student distribution may be the best for the fact that it almost covered all the actual loss. But one concern for this model is that during the 80th to 90th week it changes huge for the unknown reason, besides this conern it is generally the best method for capture the real loss.

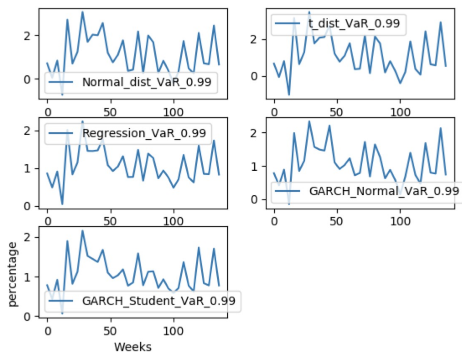
Then, by using the equation (3.8) and equation (3.9), we can plot several pictures to display how far the actual monthly losses from the VaR and $CVaR$ as shown in figure 4.18



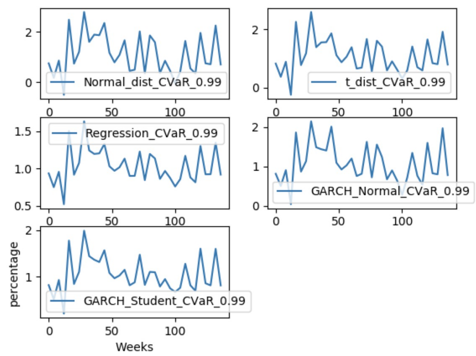
(A) $StaV_{0.95}$.



(B) $StaCV_{0.95}$.



(C) $StaV_{0.99}$.



(D) $StaCV_{0.99}$.

FIGURE 4.18: $StaV$ and $StaCV$ among different models.

Similarly to the previous work, we know when the value of Y-axis is below 0, it represents our real loss of the portfolio is bigger than the estimated VaR or CVaR we have in our models. And if the line is smooth (as shown in the Hill estimator case) that means the difference between the estimated value and the real value is big which indicates a bad model. In general, after we remove the bad models, the remaining ones are basically good enough. When at a confidence level of 0.95, StaCV performs better than StaV, but the difference between models in not clear enough, both of them have little values under zero. When at a confidence level of 0.99, we can clearly see the GARCH and Regression models perform better than others as there is seldom any value under zero which means capture all the loss situations. This result is somehow accordance to the result we get from 4.17.

4.7 Paper Trading Part

After observing the methods from previous project work, in this part, we focus on implement those methods with our holding portfolio in Interactive Broker. The holding period is form 2/26/18 to 4/19/18 with 15 stocks, 2 covered call option, 1 European put option, 2 bonds.

First, we create the QQ plot to testify if the data we use satisfy the normal distribution:

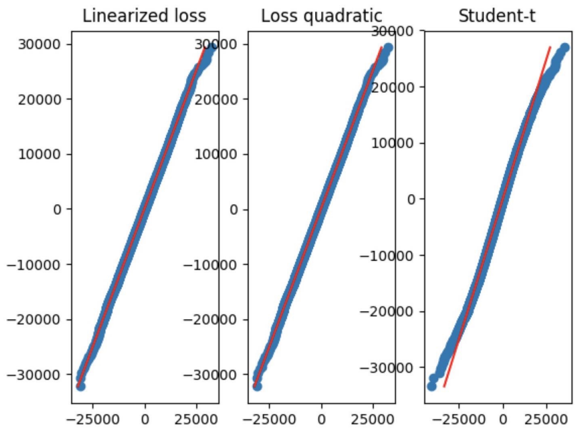


FIGURE 4.19: QQ-plot.

From 4.19, we conclude that we can use the normal distribution and t-distribution to calculate the VaR and CVaR that we are going to talk about them later.

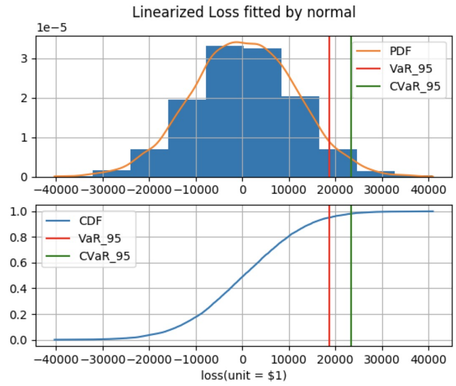
4.7.1 Stress Test

In this part, the main focus is that test our portfolio performance under some specific economy crisis time. We omit the bonds in this part, because we construct the bonds as our risk-free assets and usually we would not trade bonds before maturity time.

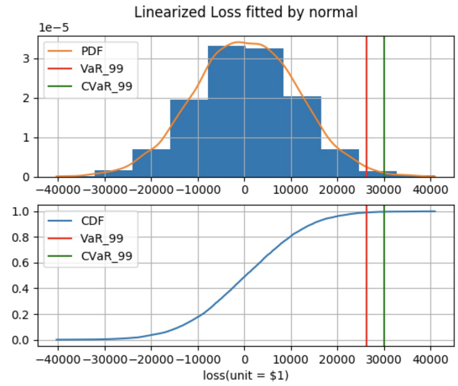
SVaR-based Analysis

In this we use 3 historical scenarios of 251 trading days which are data in 2008,2011 and 2015 to fit our model and calculate the VaR for comparison.

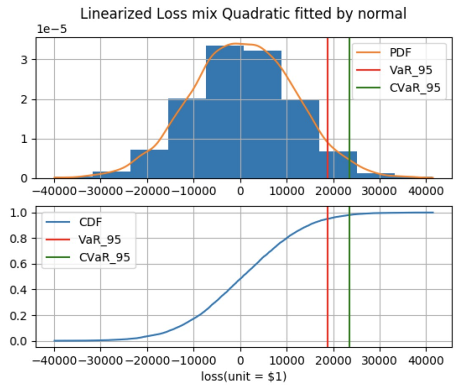
First, the following figures show us the results using data in 2008:



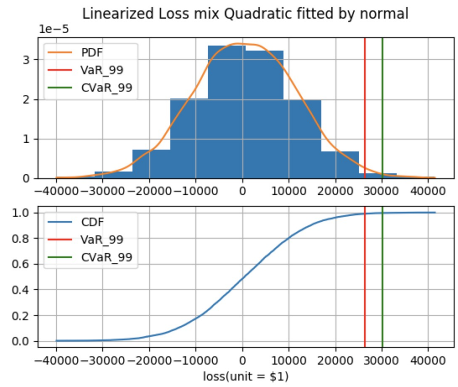
(A) $VaR_{.95}$ from linearized loss



(B) $VaR_{.99}$ from linearized loss



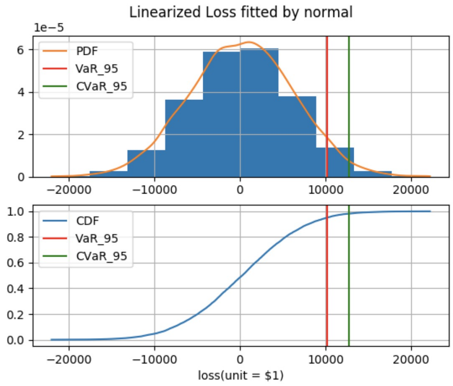
(C) $VaR_{.95}$ from quadratic loss



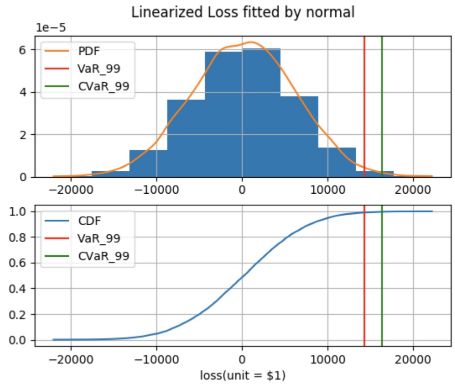
(D) $VaR_{.99}$ from quadratic loss

FIGURE 4.20: VaR and CVaR from data in 2008

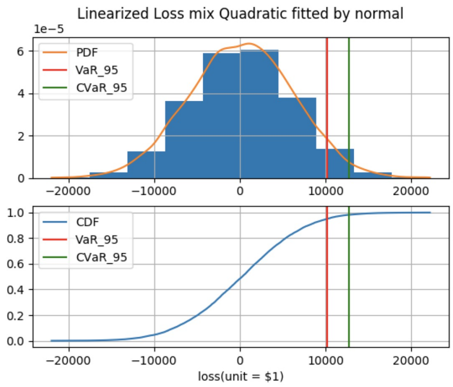
Then, we get the figures using data in 2011:



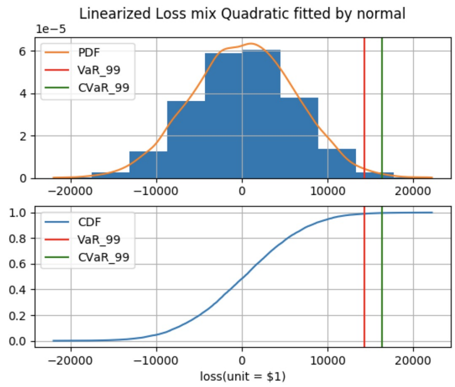
(A) $VaR_{.95}$ from linearized loss



(B) $VaR_{.99}$ from linearized loss



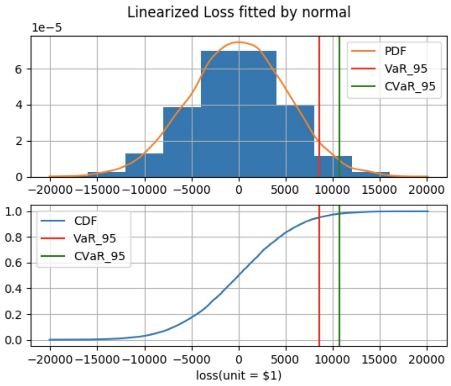
(C) $VaR_{.95}$ from quadratic loss



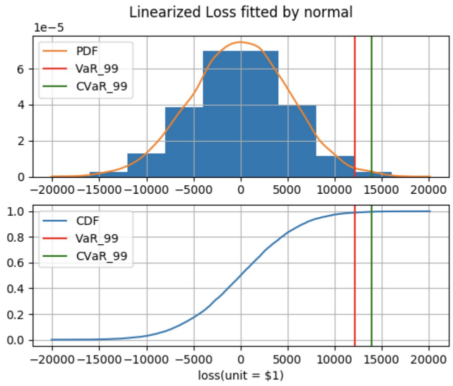
(D) $VaR_{.99}$ from quadratic loss

FIGURE 4.21: VaR and CVaR from data in 2011

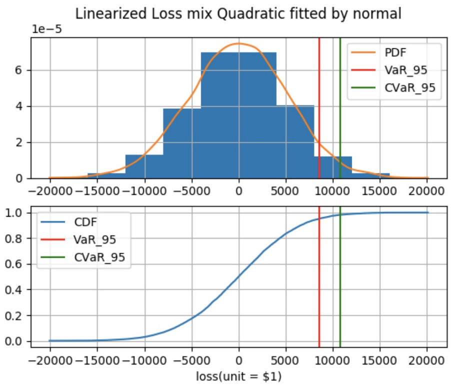
Finally,we get figures using data in 2015:



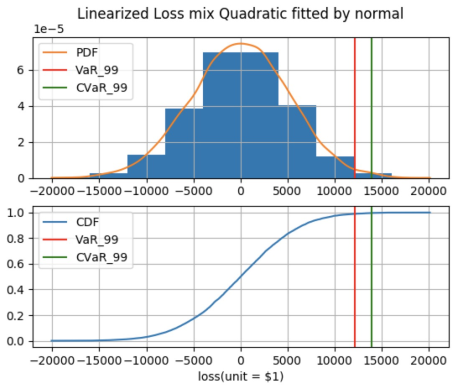
(A) $VaR_{.95}$ from linearized loss



(B) $VaR_{.99}$ from linearized loss



(C) $VaR_{.95}$ from quadratic loss



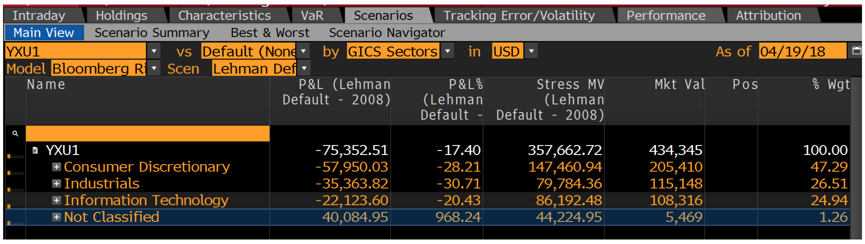
(D) $VaR_{.99}$ from quadratic loss

FIGURE 4.22: VaR and CVaR from data in 2015

Together with the figures above, under the same confidence level the linearized loss model and linearized loss mix quadratic model are basically the same. We can see under the 2008 crisis gives us the biggest VaR when under the same confidence level and 2015 crisis gives us the smallest VaR. This in fact is accordance with the result we get from Bloomberg in 4.23. That actually our portfolio is more risker under 2008 than in 2011 and 2015. The results we get from 2015 is similar to the VaR calculated by Bloomberg for recent days which is an evidence that the results we get are reasonable.

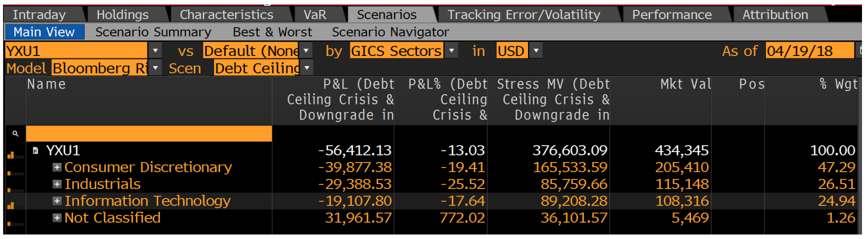
Bloomberg Scenarios

After doing the stress test with our own analysis method, we use the Bloomberg PRTU function to actually build a real portfolio at 2/26/18. With the analyze function, we are able to actually simulate our portfolio under those 3 historical time frame to see how would our portfolio probably performed under that circumstances. The three scenarios we choosed are: 2008-Lehman Default; 2011-Debt Ceiling & Downgrade; 2015- Greece Financial Crisis. The following figures show how might our portfolio performance be at those scenarios:



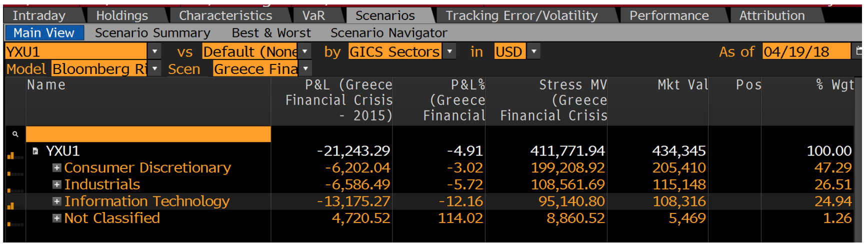
Name	P&L (Lehman Default - 2008)	P&L% (Lehman Default -	Stress MV (Lehman Default - 2008)	Mkt Val	Pos	% Wgt
YXU1	-75,352.51	-17.40	357,662.72	434,345		100.00
Consumer Discretionary	-57,950.03	-28.21	147,460.94	205,410		47.29
Industrials	-35,363.82	-30.71	79,784.36	115,148		26.51
Information Technology	-22,123.60	-20.43	86,192.48	108,316		24.94
Not Classified	40,084.95	968.24	44,224.95	5,469		1.26

(A) Simulation in 2008



Name	P&L (Debt Ceiling Crisis & Downgrade in	P&L% (Debt Ceiling Crisis &	Stress MV (Debt Ceiling Crisis & Downgrade in	Mkt Val	Pos	% Wgt
YXU1	-56,412.13	-13.03	376,603.09	434,345		100.00
Consumer Discretionary	-39,877.38	-19.41	165,533.59	205,410		47.29
Industrials	-29,388.53	-25.52	85,759.66	115,148		26.51
Information Technology	-19,107.80	-17.64	89,208.28	108,316		24.94
Not Classified	31,961.57	772.02	36,101.57	5,469		1.26

(B) Simulation in 2011



Name	P&L (Greece Financial Crisis - 2015)	P&L% (Greece Financial	Stress MV (Greece Financial Crisis	Mkt Val	Pos	% Wgt
YXU1	-21,243.29	-4.91	411,771.94	434,345		100.00
Consumer Discretionary	-6,202.04	-3.02	199,208.92	205,410		47.29
Industrials	-6,586.49	-5.72	108,561.69	115,148		26.51
Information Technology	-13,175.27	-12.16	95,140.80	108,316		24.94
Not Classified	4,720.52	114.02	8,860.52	5,469		1.26

(C) Simulation in 2015

FIGURE 4.23: Historical simulations in Bloomberg

From 4.23, we assume our standing time is 4/19/18, our portfolio value is around 434,345, Bloomberg help us classify our assets into 3 major categories. As we can see, both crisises would hurt our portfolio a lot. The 2008 financial crisis would hurt us the most, we would lose 17.4% of the total value at that time but during 2015 financial crisis, we would only lose 4.91%, that is a big difference.

4.7.2 Risk reduction

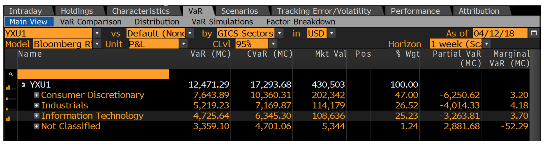
In this part, we want to reduce the risk of our portfolio by 15% due to the uncertainty of the recent market. Throughout the whole project work, we consider the VaR and CVaR as main part of our risk report, so we decide to use CVaR as our tool to solve the problem. We want to calculate the $CVaR_{.99}$ using the suitable method and then cover that value by buying the same value of the put options.

At first, we consider covered call or put option for some particularly high volatility stocks, but we find our they always act similar recently which means if something shocked happens in the market, they will all going up or down, so it's more properly dealing with the index. After consideration, we decide to buy index's put option related to S&P500 to hedge the risk.

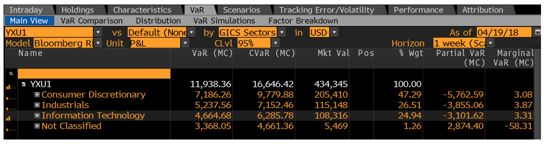
The underlying we choose is SPY (the ETF that replicates the S&P 500) which is very liquid and trades widely in markets. The options on the SPY trade in penny increments for options priced less than $3 in premium, the trading volumn is high and they have weeekly expirations.

Bloomberg Analysis

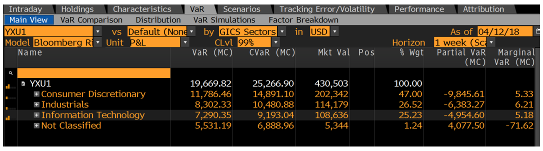
In this part, we use the portfolio analyze function in Bloomberg to test the result of our risk reduction work. we choices of our standing point is 4/12/18 and 4/19/18, one is the day before the risk reduction, one is the day we close our portfolio. We create the following figures to show the VaR at 0.95 and 0.99 confidence level:



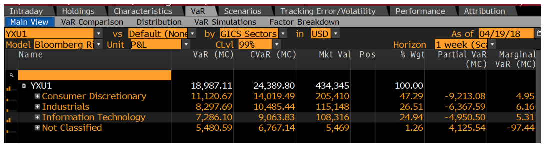
(A) $VaR_{.95}$ before reduction



(B) $VaR_{.95}$ after reduction



(C) $VaR_{.99}$ before reduction



(D) $VaR_{.99}$ after reduction

FIGURE 4.24: Comparison of VaR before and after reduction

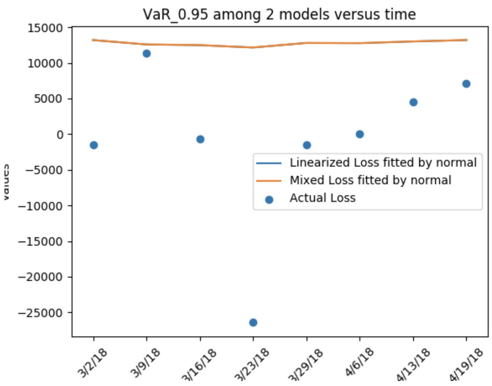
Form 4.24, we can see the $VaR_{.95}$ before is 12,471.29 and after is 11,938.36 which reduces 4.27%; the $VaR_{.99}$ before is 19,669.82 and after is 18,987.11 which reduces 3.47%.

4.7.3 Risk Reporting

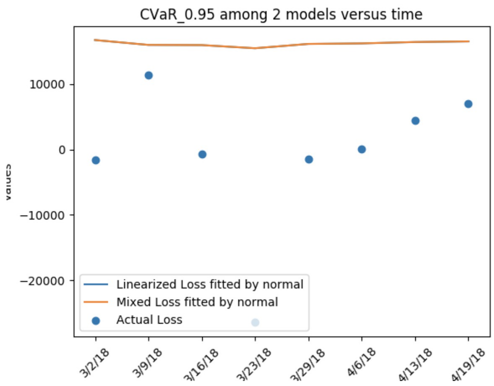
Similar to what we have done in the previous project work, here we want to analysis how our real portfoliio's risk changes over time.

Linearized Loss model and Mixing linearized & Quadratic model

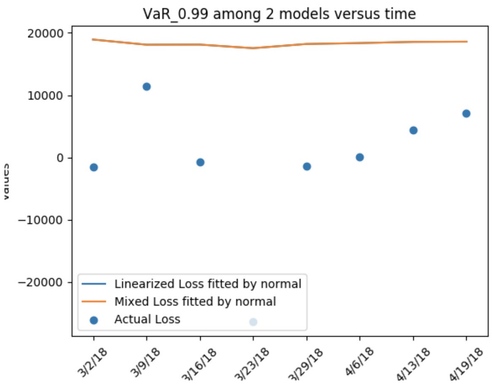
Through 8 holding weeks, we calculate our portfolio's real loss and compare to the VaR & CVaR we get from the two models.



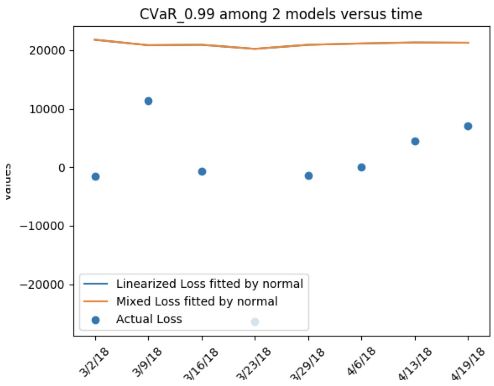
(A) $VaR_{.95}$ from two models



(B) $CVaR_{.95}$ from two models



(C) $VaR_{.99}$ from two models



(D) $CVaR_{.99}$ from two models

FIGURE 4.25: Comparison between two models

From 4.25, we can see the results of VaR and CVaR we get from two models are basically the same as it seems like only one line in the plot which are in accordance with the results we get in the stress test session. And the results we get from each model are both higher than the actual loss which means the models are efficient to capture the loss. But we find out the trend of the prediction results are not as fluctuated as the actual loss. The reason we believe is that we have short time frame, with enough holding time, we may see the trend more clearly.

Results Table among three models

Together with the two models we talked above, we also use the t-distribution wo model the overall loss of our portfolio. We create two tables showing the $VaR_{.95}$ both in a dollar version and a percentage version:

VaR_0.95	3/2/18	3/9/18	3/16/18	3/23/18	3/29/18	4/6/18	4/13/18	4/19/18
Linearized	12983.6993011	12963.7744917	12241.7278304	12078.5508829	12799.4447751	12773.0325327	12911.0813187	13217.306803
Mixed	12988.3244087	12968.7276846	12247.5561316	12084.6758203	12805.1675087	12779.2415232	12918.2726135	13224.7310776
T student	10953.2924024	1.38932530106e+26	1.02130509976e+87	10403.7332593	11721.4278403	11437.6617111	1.0197732172e+91	10947.3850783

(A) dollar version

VaR_0.95(%)	3/2/18	3/9/18	3/16/18	3/23/18	3/29/18	4/6/18	4/13/18	4/19/18
Linearized	1.32	1.28	1.24	1.23	1.26	1.28	1.30	1.33
Mixed	1.32	1.28	1.24	1.23	1.26	1.28	1.30	1.33
T student	1.10	1.18261234041e+19	1.53697582875e+36	1.04	1.17	1.13	1.02130509998e+83	1.14

(B) percentage version

FIGURE 4.26: VaR calculated by three models

From 4.26, we can see the results between the Linearized and Mixed model are really close. But the results we get from t-distribution varies too much, sometimes we get extreme large values. That is because our data set is too small to fit the t-distribution model and that is why we did not show the comparison between actual loss and t-distribution in the previous figures.

ARMA-GARCH(1,1)

In order to capture the fluctuation of the actual loss during our portfolio, we decide to try ARMA-GARCH(1,1) to fit our data. We show the comparison in the following figure:

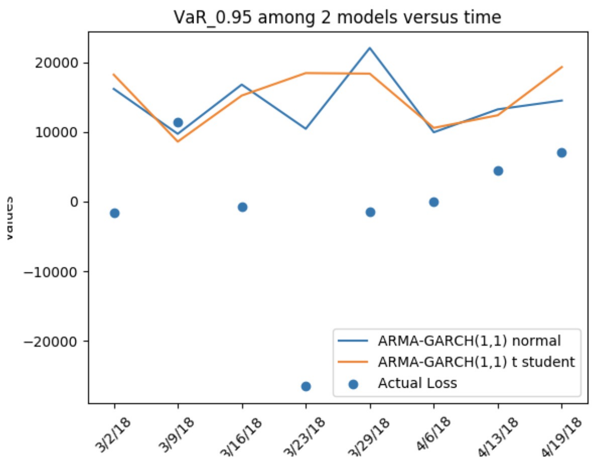


FIGURE 4.27: $VaR_{.95}$ prediction using GARCH

From 4.27, we can see by using ARMA-GARCH(1,1), we do have some fluctuation which can represent the trend from the actual loss. We have one actual loss that is slightly bigger than the VaR we get from two models, but at 0.99 confidence level we will sure capture that loss. So in all, we believe that GARCH works well in this case and if given a longer time frame we can see better results.

4.8 Conclusion

In this project, we really learn a lot hands on experience by actually conduct a portfolio in Interactive Broker, and the assets we choose including stocks, options and bonds. By the risk analysis method we choose, we now know how to properly measure the risk of the portfolio under different condition.

We think we have choosen the right model to evaluate the risk as the results we get capture the real loss we have during the holding period and throughly explain the reasons behind. In general, we believe the risk of our portfolio is not high and somehow steady on a level. As you can see, each week's $VaR_{.95}$ are around 13000 although the market experience a lot ups-and-downs.

So it's reasonable to consider covered call and put option to hedge some risk when you construct your investment portfolio. But there are a lot improvement we can make, like we can better do the risk reduction part or we can increase our portfolio holding period to better fit the model and risk reporting.

Bibliography

A. Thavaneswaran, B. Abraham (1988). *Estimation of nonlinear time series models using estimating equations*. J. Time Series Anal., 9, 99-108.

Blais, Marcel (2017). *Portfolio Valuation and Risk Management Lecture Notes*. Unpublished.

– (2018). *Market and Credit Risk Models and Management Notes*. Unpublished.

C. He, T. Terasvirta (1999). *Properties of moments of a family of GARCH processes*. Journal of Econometrics, 92, 173-192.

Cyril Coste Rapha´el Douady, Llija I. (2009). *The StressVaR: A new risk concept for superior fund allocation*. RiskData S.A., 6, 75001.

Engle Robert F., Mezrich Joseph (1996). *GARCH for Groups*. Risk, 9(8), 36-40.

Lves AR Abbott KC, Ziebarth NL (2010). *Analysis of ecological time series with ARMA(p,q) models*. Ecology. 91(3):858-71.

Nelson, Daniel B. (1991). *Conditional Heteroscedasticity in Asset Returns: A New Approach*. Econometrica, 59(2), 347-370.

R. Rabemananjara, J. M. Zakoian (1993). *Threshold Arch Models and Asymmetries in Volatility*. Journal of Applied Econometrics, Vol.8(1), 31-49.

Turner, Evan (2010). *The black-scholes model and extensions*. Unpublished.

White, H. (1980). *A Heteroskedastic-Consistent Covariance Matrix and a Direct Test fot Heteroskedasticity*. Econometrica, 48, 421-448.