Chapter 8

Risk Management

8.1 VaR and Expected Shortfall

8.1.1 Value at Risk

To compute value at risk you must specify a time horizon, d, and a percentage level $p\%$. The value at risk is the $(100-p)$-th percentile of the loss distribution at time d.

As there isn't universal agreement on how percentiles should be defined, this definition isn't completely clear. Also what some people call the 5% VaR, others call 95% VaR. So let us give a very precise definition.

Definition (Formal)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let X be a real valued random variable representing the profit and loss of our portfolio. We define*

$$VaR_\alpha(X) = \inf\{x \in \mathbb{R} : \mathbb{P}(X < -x) \leq 1 - \alpha\}$$

Note that what we call $\mathrm{VaR}_{0.05}$ is often (but not always) called the 95% value at risk by practitioners. For example the PORT function on the Bloomberg terminals uses the second convention.

In somewhat informal business language we have:

Definition (Informal)**.** *The $p\%$, d-day Value at Risk of a position is the maximum amount of money you lose over an d day period in the $p\%$ best case scenarios.*

The word "maximum" hints that this is informal since a mathematician would talk about a supremum.

Statistically this will be indistinguishable from the minimum amount of money you lose over an d day period in the $(100-p)\%$ worst case scenarios.

For the purposes of the FM06 exam, an informal definition is fine.

8.1.2 Expected Shortfall

The Basel II accord required banks to measure their risk using Value at Risk. As we will discuss later, there are some problems with value at risk. A combination of the criticism from the mathematical community and need to respond to the 2008 financial crisis lead to the Basel III accord requiring banks to measure their risk using Expected Shortfall.

Let us give an intuitive, informal definition first.

Definition (Informal)**.** *The p% d-day Expected Shortfall is the expected loss in the worst p% of cases.*

It isn't immediately obvious how this can be related to the official mathematical definition which we now give. Again $\mathrm{ES}_{0.05}$ corresponds to the business notion of 95% Expected Shortfall.

Definition (Formal)**.** $ES_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha VaR_\gamma(X)\mathrm{d}\gamma$

In the exam, I am happy for you to use the informal definition. But, out of interest let us see how the two definitions are related.

To relate the two, let us write F_X for the cumulative distribution function of X and let us assume that X has is a continuous distribution so has a density function p_x. These are, of course related by:

$$F_X(x) = \int_{-\infty}^{x} p(t)\,\mathrm{d}t.$$

By definition:

$$\mathrm{Var}_\alpha = \inf\{x \in \mathbb{R} : \mathbb{P}(X < -x) \leq 1-\alpha\}$$

Since the distribution of X is continuous, we have that $\mathbb{P}(X < -x) = P(X \leq -x) = F_X(-x)$. So we have:

$$\mathrm{Var}_\alpha = \inf\{x \in \mathbb{R} : F_X(-x) \leq 1-\alpha\}.$$

We now use the continuity of F_X to deduce that

$$\mathrm{Var}_\alpha = -F_X^{-1}(1-\alpha).$$

Returning the the formal definition of ES we have

$$\begin{aligned}\mathrm{ES}_\alpha(X) &= \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_\gamma(X)\mathrm{d}\gamma \\ &= \frac{1}{\alpha}\int_0^\alpha -F_X^{-1}(1-\gamma)\mathrm{d}\gamma\end{aligned}$$

Let us now integrate this by substitution. We make the substitution:

$$F_X(s) = 1-\gamma.$$

So

$$p_x(s)\mathrm{d}s = -\mathrm{d}\gamma.$$

Hence

$$\begin{aligned}
\mathrm{ES}_\alpha(X) &= \frac{1}{\alpha}\int_{-\infty}^{F_X^{-1}(1-\alpha)} s\,p_x(s)\,\mathrm{d}s \\
&= \frac{1}{\alpha}\int_{-\infty}^{\mathrm{VaR}_\alpha(X)} s\,p_x(s)\,\mathrm{d}s \\
&= \frac{\int_{-\infty}^{\mathrm{VaR}_\alpha(X)} s\,p_x(s)\,\mathrm{d}s}{\int_{-\infty}^{\mathrm{VaR}_\alpha(X)} p_x(s)\,\mathrm{d}s} \\
&= E(X | X \leq \mathrm{VaR}_\alpha(X))
\end{aligned}$$

The relationship to the informal definition is now clear.

Note that the informal definition is simple to compute, whereas the formal definition requires performing an expensive integral of VaR calculations. In computations we work with continuous distributions where we can safely use this final formula.

The reason mathematicians prefer the more obscure mathematical definition is simply that this definition makes equally good sense for discontinuous distributions while maintaining important properties of ES such as convexity.

8.2 Calculating VaR

There are two main ways of producing VaR figures.

- In Monte Carlo VaR, you use a probability model of your choice and use this to simulate asset returns and hence compute VaR figures.
- In Historic VaR, you use a probability model defined by historic returns and use this to compute VaR figures.

Thus historic VaR and Monte Carlo Var calculate value at risk using entirely different probability models.

In this section, we will describe these techniques and an approximation technique called parameteric VaR. Although we only discuss the computation of VaR, the same ideas can be applied to ES.

8.3 Monte Carlo VaR

Algorithm. *Simulate a large number of price paths of the underlyings in the* $\mathbb{P}$ *measure over an d-day horizon. For each price path, compute the loss of your position on day d. Read off the* $(100 - p)\%$ *percentile of the distribution. This gives you an approximation to the* $p\%$*, d-day VaR.*

You can compute the Value at Risk of any portfolio. We will illustrate the idea of value at risk by showing how to compute the risk for a portfolio consisting of a single call option with a known strike and maturity.

To compute Value at Risk need to be able to compute the price of the portfolio at the start time and in each scenario. We will use the Black-Scholes formula to compute the price, so will need to assume that there is a known volatility σ we can use to compute the prices at the start and end of our simulation.

An implementation of the code to compute the VaR by Monte Carlo for a call option with strike K and maturity T is given below. `percent` contains the desired confidence level for the VaR calculation. `days` contains the number of days in the calculation.

```
function [ var ] = monteCarloVar( ...
    percent,days, ...
    strike, maturity, ...
    spot, drift, volatility, ...
    riskFreeRate, ...
    nPaths)

t = days/365;
startPrice = blackScholesCallPrice( ...
    strike, maturity, ...
    spot,riskFreeRate, volatility ); % price portfolio at start
stockPrices = generateBSPaths( ...
    t, spot, drift, volatility,...
    nPaths, 1);
endPrices = blackScholesCallPrice( strike, maturity-t, ...
    stockPrices,riskFreeRate,volatility ); % price portfolio
profits = endPrices - startPrice;
var = prctile( -profits, 100-percent );

end
```

As can be seen, all of the real work is done by the `generateBSPaths` function which simulates the stock price. Note that to price more general portfolios, you should just replace the code marked with the comment "price portfolio".

8.3.1 A confidence interval for VaR

Suppose we have computed the $p\%$ d-day VaR using m samples. We would like to compute a $q\%$ confidence interval on this figure.

To do this we let $\alpha = p/100$, so $0 \leq \alpha < 1$. Similarly we let $\beta = q/100$. We now define percentage values r and s by the formulae:

$$\frac{r}{100} = \alpha + \sqrt{\frac{\alpha(1-\alpha)}{m}} N^{-1}\left(\frac{1-\beta}{2}\right)$$

$$\frac{s}{100} = \alpha - \sqrt{\frac{\alpha(1-\alpha)}{m}} N^{-1}\left(\frac{1-\beta}{2}\right)$$

With $q\%$ confidence, the $p\%$ VaR figure lies between the $r\%$ Var estimate and the $s\%$ VaR estimate.

This result is simply the standard confidence interval for a percentile and can be found in all good statistic textbooks.

8.3.2 Monte Carlo Var Summary

One calculates value at risk by simulating price changes and reading off the desired percentile.

On the plus side, this is highly flexible. You can use any pricing model and compute VaR for any security that you know how to price. It is intuitive and easy to program assuming you can already simulate your stochastic model and price your securities.

On the down side, it is slow. We may have to run tens of thousands of scenarios and then price the security in each scenario. This may require an additional Monte Carlo simulation if the security is exotic. Thus we can easily need to generate hundreds of millions of scenarios in total.

Another criticism is that Monte Carlo VaR is subjective. It depends heavily upon the choice of model used to generate prices. As a result, VaR estimates are only as good as the model used. If your model doesn't have fat tails, you will underestimate risk.

The RiskMetrics approach of using a standardized model helps somewhat with this subjectivity, although their approach of using a log normal model and an exponentially weighted moving average cannot account for fat tails. For this reason Basel III recommends more complicated models and fitting approaches such as GARCH models.

8.4 Historic VaR

As we have discussed, Monte Carlo VaR relies on having a good model of the distribution of risk factors. Unfortunately there is no general agreement on what the right model is.

The idea of historic VaR is to use historic log-returns and assume that a similar pattern will be repeated. This approach is called historic VaR.

Definition. *The log-return of an asset from time t_1 to time t_2 is*

$$r_{t_1,t_2} = \log\left(\frac{P_{t_2}}{P_{t_1}}\right)$$

where P_t is the price of the asset at time t.

Definition. *The relative-return of an asset from time t_1 to time t_2 is*

$$\frac{P_{t_2} - P_{t_1}}{P_{t_1}}.$$

Warning: the log-return of an asset is NOT the log of the relative-return of an asset.

Algorithm. *Suppose that we have $m + 1$ days of historical log-returns for N risk factors. $r_i^{(a)}$ is the log-return of risk factor a on day i. For each day i in our sample, generate a simulated d-day log-return by the formula:*

$$r_i^{(a)}\sqrt{d}$$

Use these simulated log-returns to simulate the values of our risk factor on day d. Compute the historic VaR by reading off the desired percentile.

The interesting feature of this algorithm is that it uses the familiar scaling properties of volatility to scale up 1-day changes to d-day changes. We are assuming the volatilty of all our risk factors scales in the same way as is familiar for stock prices.

The reason we need to scale up one day changes to d-day changes is that we are unlikely to have much historic data. If we want to compute 30-day VaR figures, then we wouldn't have much data available if we insisted on looking only at the historic changes in a stock price every 30 days.

If you can compute Monte Carlo VaR and have access to historic data, computing historic VaR is easy, as illustrated in the code example below.

```
function [ var ] = historicVar( ...
    percent,days, ...
    strike, maturity, ...
    spot, historicPrices, ...
    riskFreeRate, volatility)

dailyLogReturns = diff( log( historicPrices ));
dDayLogReturns = dailyLogReturns .* sqrt( days );
stockPrices = exp(dDayLogReturns) * spot;

t = days/365;
startPrice = blackScholesCallPrice( ...
    strike, maturity, ...
    spot,riskFreeRate, volatility ); % price portfolio at start
endPrices = blackScholesCallPrice( strike, maturity-t, ...
    stockPrices,riskFreeRate,volatility ); % price portfolio
profits = endPrices - startPrice;
var = prctile( -profits, 100-percent );

end
```

We have simply replaced the code that simulated the stock price with code that simulated stock prices by assuming that the historic changes will recur with suitable scaling. We highlight the key lines below:

```
dailyLogReturns = diff( log( historicPrices ));
dDayLogReturns = dailyLogReturns .* sqrt( days );
stockPrices = exp(dDayLogReturns) * spot;
```

Thus the vector `stockPrices` will contain a vector of simulated stock prices which exhibit the same log-return distribution as was found historically.

8.4.1 Historic Var Summary

Historic VaR has many good points. It is less subjective than Monte Carlo VaR. It may show fat tails behaviour if the historic data does. It is as simple to calculate as Monte Carlo Var. It is quick to calculate because the number of available scenarios will always be small.

Against this, historic VaR relies on the assumption that the past will predict the future. Also only quite crude calculations possible due to lack of historic data.

8.5 Parametric VaR

Monte Carlo VaR is time consuming to calculate. However, risk managers are keen to know how much risk they are taking on an intra-day basis so that they can quickly take advantage of any opportunities that arise in the market without inadvertently breaking their risk limits. Thus there is real commercial advantage in being able to calculate risk figures rapidly. For this reason, numerous techniques have been developed for approximating risk figures. We will discuss one simple approach called parametric VaR by RiskMetrics and which is also often called "delta normal VaR".

We let V be the price of a security that depends upon risk factors

$$P^{(1)}, P^{(2)}, \dots, P^{(n)}$$

We define $p^{(a)} = \log P^{(a)}$ and define

$$\delta^{(a)} = \frac{\partial V}{\partial p^{(a)}} = P^{(a)} \frac{\partial V}{\partial P^{(a)}}.$$

The value $\delta^{(a)}$ should be thought of as the sensitivity of the security price to the log-return of $P^{(a)}$.

If we assume that V depends smoothly upon the price of the risk factors, we will have the following Taylor series for V.

$$\begin{aligned} V(t, p^{(1)}, \dots, p^{(n)}) &= V_0 + \frac{\partial V}{\partial t} t + \sum_{i=1}^{n} \frac{\partial V}{\partial p^{(i)}} r^{(i)} + \dots \\ &= V_0 + \frac{\partial V}{\partial t} t + \sum_{i=1}^{n} \delta^{(i)} r^{(i)} + \dots \end{aligned}$$

Here $r^{(i)}$ is the log return of risk factor i.

We expect the risk factors to scale in a manner that is proportional to $\sqrt{t}$ so, the $\frac{\partial V}{\partial t}$ will be negligible compared to the terms involving $r^{(i)}$.

Neglecting higher order terms we have

$$V(t, p^{(1)}, \ldots, p^{(n)}) \approx V_0 + \sum_{i=1}^{n} \delta^{(i)} r^{(i)}$$

Hence if we assume that the log-returns are normally distributed with mean 0 and covariance matrix Σ, we can read off the mean and standard deviation of V in terms of Σ the covariance matrix of the $r^{(i)}$.

We deduce that, to order $\sqrt{d}$, the distribution for the change in V over a d-day period is normally distributed with mean 0 and standard deviation:

$$\sqrt{\frac{d\delta^T \Sigma \delta}{365}}$$

Here δ is the vector with components $\delta^{(a)}$.

Therefore the $p\%$ d-day VaR can be approximated by

$$N^{-1}\left(\frac{100-p}{100}\right)\sqrt{\frac{d\delta^T \Sigma \delta}{365}}$$

Algorithm (Parametric VaR)**.** *Compute the sensitivities of V*

$$\delta^{(a)} = P^{(a)} \frac{\partial V}{\partial P^{(a)}}$$

Note that in the case where P is a stock, S this is $S\Delta$. Use the formula

$$\text{VaR} \approx N^{-1}\left(\frac{100-p}{100}\right)\sqrt{\frac{d\delta^T \Sigma \delta}{365}}$$

8.6 Axiomatic Theory

We will call any $\rho : L^\infty(\Omega; \mathbb{R}) \to \mathbb{R}$ a *risk figure* although this terminology is non-standard. A risk figure may satisfy some of the following properties

1. Montonicity: if portfolio A can never out perform portfolio B, then portfolio A is riskier than portfolio B, $\rho(A) \geq \rho(B)$.

2. Translation invariance: If you add c units of cash to your portfolio, the risk decrease by the same amount. $\rho(A + c) = \rho(A) - c$.

3. Convexity: $\rho(\lambda A + (1-\lambda)B) \leq \lambda\rho(A) + (1-\lambda)\rho(B)$ if $\lambda \in [0, 1]$.

4. Positive Homogeneity: $\rho(\lambda A) = \lambda\rho(A)$ if $\lambda \in \mathbb{R}^+$.

5. Sub-Additivity: $\rho(\lambda A + B) \leq \rho(A) + \rho(B)$.

Definition. *A convex monetary risk measure is a risk figure which satisfies Monotonicity, Translation Invariance and Convexity. [3]*

Definition. *A coherent risk measure is a risk figure which satisfies Monotonicity, Translation Invariance, Convexity and Positive Homogeneity. [2] (This implies convexity)*

It is argued that a risk measure should certainly be convex because most people agree diversification reduces risk and convex combinations of assets are a form of diversification. It is harder (in my view) to argue that a risk measure should be coherent, though Artzner et al. do attempt to justify their axioms in their paper. I find the positive homogeneity axiom hard to accept.

Example 1: Value at Risk is not a convex monetary risk measure or a coherent risk measure.

Proof. Let X and Y be independent stocks. We sell digital put options on X and Y with maturity in d days and strikes chosen such that the probability of them ending in the money is only 4%. We sell them for their current market prices c_X and c_Y. The 5% d-day value at risk of being short a put on X is $-c_X$ since we only pay out 4% of the time and make c_X profit otherwise. The 5% d-day value at risk of being short a put on Y is $-c_Y$ since we only pay out 4% of the time. The probability of one or other of the options ending up in the money is $1-(0.96)(0.96) = 0.0784 > 0.05$. Therefore the 5% d-day value at risk of the portfolio is $-c_X - c_Y + 1$ which is greater than $-c_X - c_Y$. Therefore VaR is not sub-additive. □

Example 2: Conditional Value at Risk is a coherent risk measure. The proof of sub-additivity is a little challenging, all the other properties are obvious. You can find a proof in Follmer and Schied [4].

Example 3: If u is a concave increasing function then the expected disutility $E(-u(X))$ is monotone and convex but not translation invariant (unless $u(x) = x$). So expected utilities are not convex monetary risk measures.

Example 4: If u is as above, with $u(0) = 0$ then the cash needed to make a position acceptable is given by

$$\rho_u(X) = \inf\{c : E(u(X+c)) = 0\}.$$

ρ_u is a convex monetary risk measure.

Example 5: If u is the exponential utility function with parameter $\lambda > 0$

$$u(x) = \frac{1 - e^{-\lambda x}}{\lambda}$$

then the associated risk measure is called the entropic risk measure

$$\rho(X) = \frac{1}{\lambda} \log(E(e^{-\lambda(X)})).$$

This is a convex, but not coherent, risk measure.

The axiomatic theory suggests that VaR is not a good way of measuring risk because it does not always encourage diversification. CVaR and the entropic risk measures both might be good measures according to axiomatic theory, although if you believe positive homogeneity is essential you should rule out the entropic risk measure.

In my own work [1], I have considered an alternative perspective to the axiomatic theory which is whether risk constraints are effective on rogue traders. One way to model a rogue trader is as an individual who is trying to maximize

$$E(u(X^+))$$

where X is the profit (or loss) of their position. Because they are only interested in the profit X^+ and not any potential loss, these traders may take very risky strategies unless they are constrained. We showed that in the Black-Scholes model, whatever value at risk, expected shortfall and cost constraints are imposed

$$\sup(E(u(X^+))) = \sup u$$

where X ranges over the possible portfolio profit distributions. In other words, the trader is just as happy as they would be without the value at risk and expected shortfall constraints. We interpret this as saying that Value at Risk and Expected Shortfall do not curb the behaviour of rogue traders.

For this reason my personal view is that Expected Shortfall is not a good way of measuring risk. Note that this is a controversial viewpoint. The study of risk measurement is a complex, unresolved issue where there is still scope for good research.

8.7 Summary

VaR and ES are popular risk figures in the industry. VaR was required by Basel II, ES by Basel III.

We have seen how to calculate Monte Carlo VaR figures, Historic VaR figures and Parametric VaR figures. Parametric Var and Monte Carlo VaR provide estimates for the VaR in a $\mathbb{P}$ measure model chosen by the risk modeller. Historic VaR calculates VaR using a model based on historic data.

There is a mathematical theory of convex monetary risk measures. According to this theory, VaR can be criticized because it is not convex. This implies that VaR may discourage diversification. Expected Shortfall, however, is convex. Nevertheless Expected Shortfall is not always effective in constraining rogue traders.

Many other risk figures exist, for example the entropic risk measure, but they have not been widely taken up by the industry.

Bibliography

[1] J. Armstrong and D. Brigo. Optimizing s-shaped utility and implications for risk management. https://arxiv.org/abs/1711.00443, 2017.

[2] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.

[3] Hans Föllmer and Alexander Schied. Convex measures of risk and trading constraints. *Finance and stochastics*, 6(4):429–447, 2002.

[4] Hans Föllmer and Alexander Schied. *Stochastic finance: an introduction in discrete time*. Walter de Gruyter, 2011.