Risk Management

Leopoldo Catania
Aarhus University and CREATES
leopoldo.catania@econ.au.dk

Need for risk management

- There are many different types of risk
  - *Market Risk*: Due to changes in prices;
  - *Credit Risk*: Counter party does not meet contractual obligations;
  - *Liquidity Risk*: Extra cost of liquidating a position;
- It is crucial to develop tools to deal with these sources of risk.
- Basel Committee on Bank Supervision (BCBS) imposes capital requirements to cover these risk.
- In 1996 agreement on market risk, and introduction of the *Value-at-risk* (VaR).
- Nowadays, VaR and *Expected Shortfall* (ES) are widely used because they can be applied to all securities.

VaR

- The value at risk is defined with respect to:
  - Time horizon: τ
  - Confidence level: $1 - \alpha$
- The *VaR* is a bound such that the loss over the horizon is less than this bound with probability equal to the confidence coefficient.
- In other words, the *VaR* is the minimum potential loss that the portfolio can suffer in the $\alpha\%$ worst cases, over the period τ.
- Statistically speaking, the VaR is a quantile of the return(or loss) distribution.
- For example, if the horizon is one week, and with $\alpha = 1\%$ the *VaR* is \$5 millions, this means that there is a 1% chance of a loss exceeding \$5 millions over the next week.

VaR

Given $\mathcal{L}_{t+\tau}$ the loss over the holding period τ, the $VaR(\alpha)$ at time t is the α-th upper quantile of $\mathcal{L}_{t+\tau}$. For continuous, loss distribution, the $VaR(\alpha)$ solves

$$Pr(\mathcal{L}_{t+\tau} \geq VaR_t(\alpha)) = \alpha$$

or alternatively

$$Pr(\mathcal{R}_{t+\tau} < VaR_t(\alpha)) = \alpha$$

where $\mathcal{R} = -\mathcal{L}$ is the revenue, defined as

$$\mathcal{R}_{t+\tau} = \frac{\Delta W_{t+\tau}}{W_t} \tag{1}$$

where W_t is the value of the portfolio. Hence, the VaR is

$$VaR_{\alpha,t+\tau|t} = F^{-1}_{t+\tau|t}(\alpha) \tag{2}$$

where $F_{t+\tau|t}(x) = Pr(R_{t+\tau} \leq x|\mathcal{F}_t)$.

VaR: Properties

Artzner et. al (1997,1999) list the properties that any risk measure, $\rho(\cdot)$ should have to be *coherent*:

- Monotonicity: if $X \leq Y$, then $\rho(X) \geq \rho(Y)$
- Homogeneity: if $\kappa \geq 0$ then $\rho(\kappa X) = \kappa \rho(X)$
- Translation Invariance: if F is a risk-free asset with return r_f, then $\rho(X + F) = \rho(X) - r_f$
- Sub-additivity: $\rho(X + Y) \leq \rho(X) + \rho(Y)$

The VaR is not a coherent risk measure because it does not respect sub-additivity. As a consequence it discourages diversification.
Note that here X and Y are loss distributions (negative of the returns)

Example

- Company sells a bond for \$1000 with $\tau = 1$ year and rate 5%
- If the bank defaults, the entire \$1000 is lost. The probability of default is $p = 4\%$.
- The loss function is (use the normal distribution to make the loss function continuous)

$$\mathcal{L} = (1 - p) \cdot N(-50, 1) + p \cdot N(1000, 1)$$

- Suppose another independent company selling a bond with the same loss function
- Suppose $P1$, made by 2 bonds of company 1, its loss function is

$$\mathcal{L} = 0.04 \cdot \Phi(x; 2000, 4) + 0.96 \cdot \Phi(x; -100, 4)$$

- Suppose $P2$, made by 1 bond of each company, its loss function is

$$\mathcal{L} = 0.04^2 \cdot \Phi(x; 2000, 2) + 2 \cdot (0.96) \cdot (0.04) \cdot \Phi(x; 950, 2) + 0.96^2 \cdot \Phi(x; -100, 2)$$

- $VaR(0.05)_{P1} = -95.38$ and $VaR(0.05)_{P2} = 949.53$. This seems to tells us that Portfolio 1 is much less risky than portfolio 2. Is it true?
- Result depends on the choice of α.

Expected Shortfall

- The VaR is not informative on the amount of the loss over the threshold.
- Basak and Shapiro (2001): the VaR disregards the risk of extreme large losses, i.e. large losses behind the confidence level.
- The VaR is not sub-additive.
- A newer risk measure is the *Expected Shortfall*, or ES.
- ES is defined as the expected loss given that the loss exceeds the VaR. Formally

$$ES(\alpha) = \frac{\int_{\alpha}^{1} VaR(u)du}{1-\alpha}$$

  which is the average of $VaR(u)$ over all u that are less or equal to α

$$ES(\alpha) = E[\mathcal{L}|\mathcal{L} \geq VaR(\alpha)]$$

- The ES can also be used for portfolio allocation.

VaR with Gaussian returns and constant parameters

- Suppose that the return on a stock is normally distributed with yearly mean μ and variance σ^2. Suppose that we purchase \$ 100,000 of that stock, what is the VaR for $\tau = 1$ year?
- The distribution of our position is Gaussian with mean $\mu_L = 100,000 \times \mu$ and standard deviation $\sigma_L = \sigma \times 100,000$.
- Therefore, the VaR is

$$\widehat{VaR}_t = \mu_L + \sigma_L z_\alpha$$

  where z_α is the α-th quantile of a normal distribution.

Modeling portfolio returns

- For the computation of the VaR and ES we need:
  - the probability, α
  - the horizon of the investment, τ
  - the value of the portfolio at t, W_t
  - the *cdf* of the portfolio return
- *Choosing a model*: for the conditional density of returns.
- Note of caution: *Aggregation*. We need the distribution over the period τ!

Estimating the VaR

Estimates of VaR based on historic data on prices. Assumptions:

- Stationarity of returns
- Independence of returns (must be relaxed)

Two general approaches:

- Non-parametric
- Parametric

Non-parametric estimation of VaR under independence

- Suppose that we want a confidence coefficient $1-\alpha$ for the risk measures.
- We estimate the α-th quantile of the return distribution;
- This is estimated as the α quantile of the historical sample distribution of returns,
- The *VaR* estimated non parametrically is

$$\widehat{VaR}_t^{np}(\alpha) = -W_t \times \widehat{q}(\alpha)$$

  where $W_t(\omega_t)$ is the size of the current position and $\widehat{q}(\cdot)$ is the estimated quantile.

- The estimate of the ES is

$$\widehat{ES}_t^{np}(\alpha) = -W_t \times \frac{\sum_{i=1}^{T} \tilde{r}_i}{M}$$

  where $\tilde{r}_i = r_i \times I(r_i < \widehat{q}(\alpha))$, and $M = \sum_{i=1}^{T} I(r_i < \widehat{q}(\alpha))$.

Historical simulations in practice

- Simplest and fastest way of computing VaR and ES
- Chose window size, N.
- Consider the $T - N + 1$ overlapping sub-samples $\{r_1, ..., r_N\}, ..., \{r_{T-N+1}, ..., r_T\}$.
- Each sub-sample is used to approximate the cdf of the data.
- For the VaR, sort each sub-sample for generic time t, as $\{\tilde{r}_{t-N+1}, ..., \tilde{r}_t\}$.
- Chose the $\lfloor \alpha N \rfloor$-th order statistic, then the VaR is

$$\widehat{VaR}_t^{hs}(\alpha) = -W_t \times \tilde{r}_{\alpha N, t}$$

- The expected shortfall is

$$\widehat{ES}_t^{hs}(\alpha) = -W_t \times \frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \tilde{r}_{i,t}$$

- VaR and ES obtained with this method do not vary often enough.

Estimating VaR and ES in a time-varying framework

- The assumption of independence of daily returns is too restrictive in practice.
- Daily returns display a small degree of autocorrelation
- But a great amount of volatility clustering

The dynamics in the volatility can be modelled as

- Historical simulation approach (estimates based on rolling windows)
- Semi-parametric (EVT)
- Parametric approach (ARMA-GARCH, JP-Morgan)

Filtered Historical Simulations

- FHR refers to an hybrid mechanism.
- It relies on a simple resampling scheme,
- The term *filtered* refers to the fact that the quantiles are based the set of shocks, $\hat{z}_t = r_t / \hat{\sigma}_{t|t-1}$, which are returns filtered by the GARCH model.
- The percentile, $\widehat{q}_z(\alpha)$, is calculated from the set of historical shocks, $\{\hat{z}_1, \hat{z}_2, \dots\}$.
- The VaR is

$$\widehat{VaR}^{fhs}_{t+1|t}(\alpha) = -W_t \times \widehat{\sigma}_{t+1|t} \cdot \widehat{q}_z(\alpha)$$

- The Expected Shortfall for the one-day horizon can be calculated as

$$\widehat{ES}^{fhs}_{t+1|t}(\alpha) = -W_t \times \widehat{\sigma}_{t+1|t} \frac{1}{\lfloor \alpha N \rfloor} \sum_{i=1}^{\lfloor \alpha N \rfloor} \tilde{\hat{z}}_{i,t}$$

GARCH-EVT model

- Estimate a GARCH model on the return series, r_t, for $t = 1, ..., T$,
- Compute the standardized residuals $\hat{z}_t = r_t / \hat{\sigma}_{t|t-1}$,
- Estimate by ML the parameters, ξ, ψ, of the generalized Pareto distribution based on the N_u exceedances below a threshold u.
- Given ξ and ψ, the α quantile of $\hat{z}_t$ is given by inverting the cdf of the exceedances, i.e.

$$\hat{q}_z(\alpha) = \begin{cases} u + \frac{-\hat{\psi}}{\hat{\xi}} \left(\left(\frac{T}{N_u} \alpha \right)^{-\hat{\xi}} - 1 \right), & \text{if } \xi \neq 0 \\ u + \hat{\psi} \log \left(\frac{T}{N_u} \alpha \right), & \text{if } \xi = 0 \end{cases}$$

- So the VaR is

$$\widehat{VaR}^{evt}_{t+1|t}(\alpha) = -W_t(\omega_t) \times \widehat{\sigma}_{t+1|t} \widehat{q}_z(\alpha)$$

- ...and the ES is

$$\widehat{ES}^{evt}_{t+1|t}(\alpha) = -W_t(\omega_t) \left(\frac{\widehat{\sigma}_{t+1|t} \widehat{q}_z(\alpha)}{1 - \hat{\xi}} - \frac{\hat{\psi} - \hat{\xi} u}{1 - \hat{\xi}} \right)$$

Parametric estimation of VaR under independence

- Parametric estimation is based on assumptions on the distribution of returns
- For example we can assume returns (or the standardized returns) to be Gaussian or Student's t distributed.
- Let $F(r|\theta)$ be a family of distributions used to model the return distribution and suppose $\widehat{\theta}$ is an estimate of θ, then the VaR is

$$\widehat{VaR}_t^{par}(\alpha) = -W_t(\omega_t) \times F^{-1}(\alpha|\widehat{\theta})$$

- $F(r|\theta)$ gives a full description of the probability of the returns for any α.

Parametric estimation of ES under independence

- The estimate of the ES is

$$\widehat{ES}_t^{par}(\alpha) = -\frac{W_t(\omega_t)}{\alpha} \times \int_{-\infty}^{F^{-1}(\alpha|\widehat{\theta})} x \cdot f(x|\widehat{\theta})dx$$

- Computing this integral may be complicated for non-standard CDFs.
- If returns are Student's t distributed with mean, μ, scale λ and ν degrees of freedom, then

$$\widehat{ES}_t^{t-stud}(\alpha) = W_t(\omega_t) \times \left\{ -\mu + \lambda \cdot \left(\frac{f_\nu[F_\nu^{-1}(\alpha)]}{\alpha} \left[\frac{\nu + [F_\nu^{-1}(\alpha)]^2}{\alpha} \right] \right) \right\}$$

- Under Gaussianity

$$\widehat{ES}_t^{norm}(\alpha) = W_t(\omega_t) \times \left\{ -\mu + \sigma \cdot \left(\frac{\phi([\Phi^{-1}(\alpha)])}{\alpha} \right) \right\}$$

where $\phi(\cdot)$ is the density of the standard Gaussian distribution and $\Phi^{-1}(\alpha)$ is the iverse cdf (quantile function) of the standard Gaussian distribution.

Estimating VaR: time-varying volatility

- Assume that $\tau = 1$ and we have T returns that we need to estimate VaR and ES for next period $T + 1$.
- Let $\widehat{\mu}_{t+1|t}$ and $\widehat{\sigma}_{t+1|t}$ the conditional mean and volatility of tomorrow's return
- Under Gaussianity, the $VaR_t(\alpha)$ is

$$\widehat{VaR}_{t+1|t}(\alpha) = -W_t(\omega_t) \times \left\{\widehat{\mu}_{t+1|t} + \widehat{\sigma}_{t+1|t}\Phi_\alpha^{-1}\right\}$$

- Under t-Student's distribution

$$\widehat{VaR}_{t+1|t}(\alpha) = -W_t(\omega_t) \times \left\{\widehat{\mu}_{t+1|t} + \widehat{\lambda}_{t+1|t} q_\alpha(\widehat{\nu})\right\}$$

where $\widehat{\lambda}_{t+1|t} = \sqrt{(\widehat{\nu} - 2)/\widehat{\nu}} \cdot \widehat{\sigma}_{t+1|t}$.

Estimating VaR: Riskmetrics approach

- Very simple way to make variance a time-varying process (assume $\mu_{t|t-1} = 0$)
- The law of motion of the variance is

$$\widehat{\sigma}^2_{t+1|t} = \delta \cdot \widehat{\sigma}^2_{t|t-1} + (1-\delta) \cdot r_t^2 \tag{3}$$

  where $0 < \delta < 1$ and the usual choice is $\delta = 0.94, 0.96$.
- The initialization is $\sigma^2_{2|1} = r_1^2$.
- Under Gaussianity, the $VaR_{t+1|t}(\alpha)$, with $\alpha = 1\%$ is

$$\widehat{VaR}^{JP}_{t+1|t}(\alpha) = -W_t(\omega_t) \times -2.326 \times \widehat{\sigma}_{t+1|t}$$

- The $ES_t(\alpha)$ at $\alpha = 1\%$ is

$$\widehat{ES}^{JP}_{t+1|t}(\alpha) = W_t(\omega_t) \times \frac{\phi(-2.3226)}{0.01} \times \widehat{\sigma}_{t+1|t}$$

- If RV is available, then the law of motion of volatility can be replaced with

$$\widehat{\sigma}^2_{t+1|t} = \delta \cdot \widehat{\sigma}^2_{t|t-1} + (1-\delta) \cdot RV_t \tag{4}$$

Multi step ahead VaR

Assume that volatility is estimated by a GARCH process and returns are assumed to be conditionally Gaussian. Also assume $W_t(\omega_t) = 1$ and $\mu = 0$. We saw that

$$\widehat{VaR}_{t+1|t}(\alpha) = -\widehat{\sigma}_{t+1|t}\Phi_{\alpha}^{-1},$$

what about $\widehat{VaR}_{t+h|t}(\alpha)$ for $h > 1$? We have closed form predictions for the h-step ahead volatility $\widehat{\sigma}_{t+h|t}$ but we don't know the distribution of $Y_{t+h}|\mathcal{F}_t$. In the literature the following approximation has been proposed:

$$\widehat{VaR}_{t+h|t}(\alpha) \approx -\widehat{\sigma}_{t+h|t}\Phi_{\alpha}^{-1},$$

however, since we know that $Y_{t+h}|\mathcal{F}_t$ has tails fatter than a Gaussian random variable, this approximation can severely underestimate the true $VaR_{t+h|t}(\alpha)$. A solution is to use Monte Carlo simulation: i) simulate from $Y_{t+h}|\mathcal{F}_t$ and, ii) estimate $\widehat{VaR}_{t+h|t}(\alpha)$ as the empirical quantile of the simulated draws.

Multi Step ahead VaR: Gaussian approximation

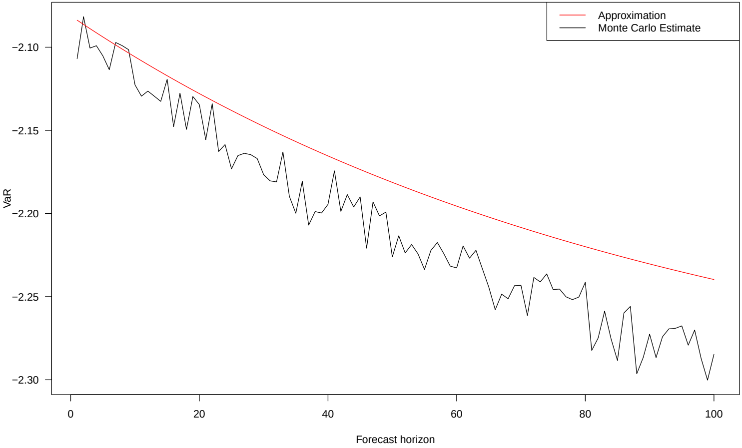

Figure: Comparison between the Gaussian approximated VaR and it's Monte Carlo estimate at $\alpha = 1\%$. See the script Lecture13.R

Back-testing VaR

- How can we compare alternative predictions of the VaR based on different econometric methods for the estimation?
- Several testing methods are generally employed
  1. Kupiec test - unconditional coverage
  2. Christoffersen test - conditional coverage
- Loss functions

Unconditional Coverage test

- Under the null hypothesis that the model is *correct*, the number of exceptions follows a binomial distribution.
- We denote $I_t(\alpha)$ the hit variable associated to the ex-post observation of a $VaR(\alpha)$ exception at time t.
- The unconditional probability of a violation must be equal to the a coverage rate

$$P(I_t(\alpha) = 1) = E(I_t(\alpha)) = \alpha$$

- Each variable $I_t(\alpha)$ has a Bernoulli distribution with probability α.

BIS Traffic Light System

Zone	Number of exceptions	Increase in scaling factor
Green Zone	0	0,00
	1	0,00
	2	0,00
	3	0,00
	4	0,00
Yellow Zone	5	0,40
	6	0,50
	7	0,65
	8	0,75
	9	0,85
Red Zone	10 or more	1,00

Note: VaR(1%, 1 day), 250 daily observations

Kupiec test

- Kupiec (1995) test attempts to determine whether the observed frequency of exceptions is consistent with the frequency of expected exceptions according to the VaR model at α level.
- The test is a LR test

$$LR_{UC} = -2\log\left[(1-\alpha)^{T-H}\alpha^{H}\right] + 2\log\left[(1-H/T)^{T-H}(H/T)^{H}\right] \to \chi^2(1)$$

where $H = \sum_{t=1}^{T} I_t(\alpha)$ denotes the total number of exceedances.

Christoffersen test

- Problem with Kupiec test: Clustering in the exceedances.
- Need to test also for independence of violations.
- Christoffersen (1998) assumes that the violation process $I_t(\alpha)$ can be represented as a Markov chain with two states:

$$\Pi = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix}$$

  where $\pi_{ij} = Pr[I_t(\alpha) = j | I_{t-1}(\alpha) = i]$, i.e. the probability of an i on day $t-1$ being followed by a j on day t.
- Under independence, $\pi_{01} = \pi_{11}$.
- The idea behind is that clustered violations represent a signal of risk model misspecification

Christoffersen test

- The corresponding LR statistic for independence is defined by:

$$
\begin{aligned}
LR_{CC} &= -2\log\left[(1-\alpha)^{T-H}\alpha^{H}\right] \\
&\quad +2\log\left[(1-\pi_{01})^{T_0-T_{01}}\pi_{01}^{T_{01}}(1-\pi_{11})^{T_1-T_{11}}\pi_{11}^{T_{11}}\right] \to \chi^2(2)
\end{aligned}
$$

where T_{ij} denotes the number of observations with j followed by i, while $\widehat{\pi}_{01} = T_{01}/T_0$ and $\widehat{\pi}_{11} = T_{11}/T_1$.

Controlling for the magnitude of the violations

- Lopez (1998) proposes to evaluate the performance of different *VaR* forecasts based on loss functions
- The loss functions that reflect the concerns of financial institutions for large deviations

$$Loss_t^f = \begin{cases} 1 + (VaR_t(\alpha) - R_t)^2 & \text{if } R_t < -VaR_t(\alpha) \\ 0 & \text{if } R_t > -VaR_t(\alpha) \end{cases}$$

- Another widely used loss function is the quantile loss function:

$$Q_t = (R_t - VaR_t(\alpha))(\alpha - \mathbb{1}(R_t < VaR_t(\alpha)))$$

- From the point of view of the individual firm, the *VaR* is an opportunity cost, so they want to minimize it

$$Loss_t^i = VaR_t$$