Risk Management 1

Lecture 1

Course Outline: 1

- This is a course on Risk Management for all MSCF students.
- It will focus on Market Risk and Credit Risk.
- It is organized with regard to the framework of GARP (Global Association of Risk Professionals), a group that administers the FRM 1 and 2 certification exams (Financial Risk Manager).
  - Note, a color-coded list of topics covered in these exams along with the learning objectives of FRM study is contained on Canvas.
  - A number of topics will be covered in much greater depth, while some will be completely omitted from the course.

Course Outline: 2

- Market Risk
  - Risk measures, coherent risk measures, and approximations
  - “Risk management in the financial industry,” Lecture on Sept 11 by Gary Lee, Global Head of Market Risk with PNC.
  - Stress testing (CCAR), Lecture, October 9 by Peter Cai, Global Head of Risk, Global Atlantic, formerly of Morgan Stanley.
  - Extreme value distributions

Course Outline: 3

- Credit Risk
  - Default modeling, reduced form models, Cox processes
  - Credit migration
  - Credit default swaps (CDS)
  - CVA/DVA (introduction - focus of Risk Management 2)
  - Economic capital, Vasicek's formula and extensions

Syllabus Essentials

- Course content
- Work required
    - Five homework assignments, two associated with practitioner presentations. One will be a team assignment involving a risk assessment of a complex portfolio with many different distinct asset classes.
    - 2 hour final exam, tentatively scheduled for October 18, 5:30-7:30.
- TA sessions (Fridays, 11:00-12:00, **if scheduled**)
- Learning Objectives
- Academic integrity and class attendance

Textbooks and Other Materials

- McNeil, Frey and Embrechts, *Quantitative Risk Management*
- Danielsson, *Financial Risk Forecasting*
- Jorion, *Value at Risk*, 3rd edition.
- Allen, *Financial Risk Management*. 2nd edition
- Glasserman, *Monte Carlo Methods in Financial Engineering*, 2003 (Chapter 9)
- Duffie and Singleton, *Credit Risk*, 2003
- Gregory, *Counterparty Credit Risk*, 2nd edition
- Jorion, *Financial Risk Manager Handbook*, 6th edition, FRM Guide.
- *Riskmetrics Technical Document*

Risk Categorization

There are many categorizations of risk factors, for example Duffie and Singleton present:

- Market risk (the risk of unexpected changes in prices or rates).
- Credit risk (the risk of changes in value associated with unexpected changes in credit quality or default of a counterparty).
- Liquidity risk (the risk that the costs of adjusting financial positions will increase substantially or that a firm will lose access to financing).
- Operational risk (the risk of fraud, systems failures, trading errors, model risk, internal organizational risks, etc.)
- Systemic risk (the risk of breakdowns in marketwide liquidity or chain reaction default)

Market Risk

- Market risk generally has 4 categories:
  1. Equity risk
  2. Interest rate risk
  3. Currency risk
  4. Commodity risk

Model Risk

Model Risk consists of:

- Violation of assumptions
- Missing factors
- Calibration errors
- Parameter uncertainty

and can lead to trading errors.

Three Fundamental Questions

- What are the relevant risk factors?
- What statistical model accurately, yet conveniently describes
  - The movements of the individual sources of risk?
  - The co-movements of multiple sources of risk?
- How does the value of a portfolio change in response to the changes in those underlying sources of risk?

The losses from individual functions or desks must be aggregate firm-wide, and the firm must have sufficient capital to sustain operations in the face of these losses.

Measuring Market Risk

- Suppose we define
  $\boldsymbol{X}$ = Vector of m market risk factors (e.g. prices, rates, etc.)
  Δ = Time Horizon (e.g. Minutes, 1 or 10 days, 1 year)
  $V(t) = f(\boldsymbol{X}, t)$ = Value of the portfolio in question at time t with current risk factor values $\boldsymbol{X}$, often referred to as a mapping of risk factors.
- Hold the portfolio fixed (no trading) during the time window $[t, t+\Delta]$. Changes in the value of the portfolio are entirely due to changes in the risk factors.
- Gain = $V(t+\Delta) - V(t) = f(\boldsymbol{X} + \Delta\boldsymbol{X}, t+\Delta) - f(\boldsymbol{X}, t)$
  = - loss over $[t, t+\Delta]$ in value over Δ.
- Loss = L = $V(t) - V(t+\Delta) = f(\boldsymbol{X}, t) - f(\boldsymbol{X} + \Delta\boldsymbol{X}, t+\Delta)$
- L is a random variable (because $\Delta\boldsymbol{X}$ is random).
- $F_L(x) = P(L \leq x)$, the c.d.f. of L.

Example: Stock Portfolio: 1

- Consider a fixed portfolio of d stocks and let λ_i be the number of shares of stock i, $1 \leq i \leq d$
- Let the logarithmic stock prices be the relevant risk factors, i.e. $X_i(t) = \log(S_i(t))$. It follows that

$$V(t) = \sum_{i=1}^{d} \lambda_i \exp(X_i(t)).$$

Thus

$$L = -(V(t+\Delta) - V(t)) = -\sum_{i=1}^{d} \lambda_i S_i(t)(\exp(Y_i(t+\Delta)) - 1),$$

where

$$Y_i(t+\Delta) = \log(S_i(t+\Delta)) - \log(S_i(t)) = X_i(t+\Delta) - X_i(t).$$

Example: Stock Portfolio: 2

- Letting $w_i(t) = \lambda_i S_i(t)/V(t)$ be the proportion of the portfolio value invested in stock i at time t, linearization of the loss becomes:

$$L = -V(t)\sum_{i=1}^{d} w_i(t)Y_i(t).$$

- If $\boldsymbol{Y}$ has mean $\boldsymbol{m}$ and covariance matrix $\boldsymbol{\Sigma}$, then the mean and variance of the linearized loss over $[t, t+\Delta]$ are given by $-V(t)\boldsymbol{w}'\boldsymbol{m}$ and $V^2(t)\boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w}$ respectively.
- See Homework #1, Problem 1.

Default-Free Bond Portfolio: 1

- Consider a portfolio of d default-free zero-coupon bonds with maturity T_i and price at time t is $p(t, T_i)$, $1 \leq i \leq d$.
- Let λ_i denote the number of bonds with maturity T_i in the portfolio.
- Note $p(t, T) = \exp(-(T - t)y(t, T))$ where $y(t, T)$ is the continuously compounded yield curve at time t.
- The value of the portfolio, $V(t)$, is given by

$$
\begin{aligned}
V(t) &= \sum_{i=1}^{d} \lambda_i p(t, T_i) \\
&= \sum_{i=1}^{d} \lambda_i \exp(-(T_i - t)y(t, T_i)).
\end{aligned}
$$

Default-Free Bond Portfolio: 2

- Let the risk factors be the changes in the yield curve,

$$X_i(t + \Delta) = y(t + \Delta, T_i) - y(t, T_i).$$

- The formula for the linearized loss becomes

$$\begin{aligned} L &= V(t) - V(t + \Delta) \\ &= \sum_{i=1}^{d} \lambda_i (p(t, T_i) - p(t + \Delta, T_i)) \\ &\approx -\sum_{i=1}^{d} \lambda_i \Delta \frac{\partial}{\partial t} p(t, T_i) \end{aligned}$$

Default-Free Bond Portfolio: 3

Now

$$
\begin{aligned}
\Delta \frac{\partial}{\partial t} p(t, T_i) &= \Delta \frac{\partial}{\partial t} e^{-(T_i - t) y(t, T_i)} \\
&\approx p(t, T_i)(\Delta y(t, T_i) - (T_i - t)(y(t + \Delta, T_i) - y(t, T_i)) \\
&= p(t, T_i)(\Delta y(t, T_i) - (T_i - t) X(t, T_i))
\end{aligned}
$$

Thus the linearized loss is approximately given by

$$
L \approx -\sum_{i=1}^{d} \lambda_i p(t, T_i)(y(t, T_i)\Delta - (T_i - t) X_{t, T_i}).
$$

Default-Free Bond Portfolio: 4

- Consider a simple example: a flat yield curve $(Y(t, T) = y(t)$, and only parallel shifts over $[t, t + \Delta]$ $(y(t + \Delta) = y(t) + \theta)$. Then $\frac{\partial}{\partial t} y(t, T_i) = y'(t)$.
- Plugging this into the linearized loss function, we find

$$\begin{aligned} L &\approx -V(t)(y(t)\Delta - \sum_{i=1}^{d} \frac{\lambda_i p(t, T_i)}{V(t)} (T_i - t)\theta) \\ &= -V(t)(y(t)\Delta - D\theta) \end{aligned}$$

  where $D = \sum_{i=1}^{d} \frac{\lambda_i p(t, T_i)}{V(t)} (T_i - t)$ is the duration of the bond portfolio.
- Use techniques like immunization to manage interest rate risk with known future payments.

Portfolio of Risky Loans

- A third example involves credit risk.
- Consider a portfolio of risky loans to d different counterparties. There are many risk factors that need to be considered. The most important are
  - Default risk - the counterparty will not be able to repay the loan.
  - Interest-rate risk - the value of the cash flows from the portfolio is reduced due to rising interest rates.
  - Losses from rising credit spreads - the credit quality of the obligor(s) decreases, hence the value of their component of the loan portfolio is diminished.
- Starting with the bond portfolio example with $y(t, T_i)$ we add a credit spread component, $c(t, T_i)$ and replace $y(t, T_i)$ by $y(t, T_i) + c(t, T_i)$.
- We will discuss default modeling later in the course.

Measuring Market Risk: 1

The distribution of L depends critically on the distribution of $\Delta \boldsymbol{S}$, the change in the risk factors over $[t, t + \Delta]$. How can the distribution be chosen? What are some of the factors?

- Risk horizon?
- Tail behavior (normal or heavy tails)?
- Dependency structure across items in portfolio (copula models)?
- What model to use for changes in prices over the risk horizon? (parametric model or non-parametric model based on recent history of behavior)
- Regular or stressed simulation? CCAR (Comprehensive Capital Assessment and Review) is stress testing mandated by U.S. Federal Reserve.

Measuring Market Risk: 2

- Simulation is most often used to estimate the distribution of L or some statistic derived from that distribution. This is often computationally intensive in that it may require:
    - Nested simulation.
    - Importance sampling (to explore the upper tail of F_L as with VaR and shortfall risk.
- In a nested simulation, the outer loop runs through different risk factor change scenarios (ΔS). For each set of risk factors specified in the outer loop, simulation may be needed to determine the value of the assets in the portfolio at time $t + \Delta$.
- One would want to use all available variance reduction techniques in the inner loop to make this valuation as fast as possible.

Value-at-Risk (VaR)

- VaR is the α quantile of F_L, the loss distribution.

$$\mathrm{VaR}_\alpha = \inf\{x | F_L(x) \geq \alpha\} = F_L^{-1}(\alpha)$$

- Remember that losses are additive, but VaR is not. Suppose we aggregate losses across m desks, $L_1, L_2, \ldots, L_m$, i.e. $L = L_1 + L_2 + \cdots + L_m$ is the total loss across all desks (firmwide loss).
- $\mathrm{VaR}_L(\alpha)$ is the α-quantile of the distribution of L.
- $\mathrm{VaR}_L(\alpha) = VaR_{L_1+\cdots+L_m}(\alpha) \neq \sum_{i=1}^m \mathrm{VaR}_{L_i}(\alpha)$.
- **VaR is NOT additive!! Don't make this mistake!**
- In general, we want a risk measure to be subadditive, to model acknowledge the benefits of diversification:

$$\mathcal{R}(L_1 + \cdots + L_m) \leq \sum_{i=1}^m \mathcal{R}(L_i).$$

Criticisms of VaR

VaR has become the standard reporting measure, mandated by regulators (Basel 2), although it is being replaced by Shortfall Risk (later). One cannot expect that any single feature or measure derived from the full loss distribution can possibly convey the entirety of the loss distribution. Any single statistic will have drawbacks. Some of the criticisms of VaR include:

- VaR pays no attention to the size of the potential losses beyond the VaR value. The VaR value could be small; however, the tail of the distribution could be very long and contain huge values.
- VaR is not a "coherent risk measure."
- Difficult to estimate extreme quantiles with any accuracy.
- Most risk measures are subject to model error.

An Alternative Risk Measure

- Another measure that can complement VaR is conditional VaR (not to be confused with credit-value-at-risk, often cVaR) (also known as shortfall risk, expected tail loss, or tail VaR.). This measure is given by $E(L|L > VaR_\alpha) = S_\alpha$ for shortfall risk.
- If simulation is used to estimate VaR(α) and S_α, one generates an independent sample of observations from the loss distribution, $L_1, \ldots, L_n$.
- Compute the order statistics: $L_{(1)} \leq L_{(2)} \leq \ldots \leq L_{(n)}$. The standard estimate of VaR(α) is $L_{\lceil n\alpha \rceil}$, the appropriate order statistic, e.g. the $L_{\lceil 990 \rceil)}$ for $n = 1,000$, and $\alpha = .99$.
- The standard estimate of S_α is given by $\frac{\sum_{i=\lceil n\alpha \rceil+1}^{n} L_{\lceil i \rceil}}{n-\lceil n\alpha \rceil}$.

Some Issues With VaR and Shortfall Risk

- VaR is explicitly mandated by Basel, and now it appears that shortfall risk will be as well.
- Both measures involve the tail of the loss distribution (VaR is a quantile, shortfall risk is the mean of the tail). Large sample sizes are needed for accurate estimation.
- VaR is not “coherent,” but shortfall risk is (discussed later).
- Ultimately, the evaluation of either measure involves choosing parameter values (such as volatilities). Therefore, each has substantial uncertainty associated with it. This leads to the need to backtest the risk measure. This will be discussed later.

Formulas for VaR_α and S_α: 1

- Suppose we consider a portfolio of λ_i shares of stock i, $1 \leq i \leq n$, where the price processes form an underlying geometric Brownian motion process with diagonal covariance matrix. Suppose that the values of the portfolio at times 0 and Δ are

$$V(0) = \sum_{i=1}^{n} \lambda_i S_{i,0}$$

$$V(\Delta) = \sum_{i=1}^{n} \lambda_i S_{i,0} \exp(\mu_i \Delta + \sigma_i \sqrt{\Delta} Z_i).$$

- Linearization (Euler) gives

$$L = -\sum_{i=1}^{n} \lambda_i S_{i,0} (\mu_i \Delta + \sigma_i \sqrt{\Delta} Z_i),$$

where $\{Z_i\}$ are i.i.d. standard normals.

Formulas for VaR_α and S_α: 2

$$1-\alpha = \int_{\mathrm{VaR}_\alpha}^{\infty} dF_L(x),$$

$$S_\alpha = \int_{\mathrm{VaR}_\alpha}^{\infty} \frac{x}{1-\alpha} dF_L(x) = \int_\alpha^1 \frac{F_L^{-1}(u)}{1-\alpha} du = \int_\alpha^1 \frac{VaR(u)}{1-\alpha} du,$$

where the last two inequalities follows from the substitution $u = F_L(x)$ and $x = F_L^{-1}(u) = VaR(u)$.

- Typical distributions used for L are the normal and the t_ν. These yield special closed form expressions for VaR_α and S_α.

Formulas for VaR_α and S_α: 3

Suppose L has a $N(0,1)$ distribution.

$$\mathrm{VaR}_\alpha = \Phi^{-1}(\alpha),$$

$$\begin{aligned}
S_\alpha &= \int_{\Phi^{-1}(\alpha)}^{\infty} x \frac{\phi(x)}{1-\alpha} dx \\
&= \int_{\Phi^{-1}(\alpha)}^{\infty} \frac{x}{1-\alpha} \frac{1}{\sqrt{2\pi}} e^{-.5x^2} dx \\
&= \ldots \\
&= \frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha}.
\end{aligned}$$

Formulas for VaR_α and S_α: 4

If L has a $N(\mu, \sigma^2)$ distribution, then

$$\text{VaR}_{\alpha} = \mu + \sigma\Phi^{-1}(\alpha),$$

and

$$S_\alpha = \mu + \frac{\sigma}{1-\alpha}\phi(\Phi^{-1}(\alpha))$$

If L has a standard t_ν distribution, then VaR is the α quantile of the distribution, or

$$\text{VaR}_\alpha = T_\nu^{-1}(\alpha).$$

$$S_\alpha = \frac{t_\nu(T_\nu^{-1}(\alpha))}{1-\alpha}\left(\frac{\nu + (T_\nu^{-1}(\alpha))^2}{\nu - 1}\right),$$

where t_ν and T_ν are the pdf and cdf respectively.

Formulas for VaR_α and S_α: 5

If L has a t-distribution with median μ and scale parameter σ

$$\text{VaR}_\alpha = \mu + \sigma T_\nu^{-1}(\alpha),$$

$$S_\alpha = \mu + \sigma \frac{t_\nu(T_\nu^{-1}(\alpha))}{1-\alpha}\left(\frac{\nu + (T_\nu^{-1}(\alpha))^2}{\nu - 1}\right),$$

Note, T_ν does not have unit variance. Rather it is $\frac{\nu}{\nu-2}$, consequently one must choose σ so that $\sigma\sqrt{\nu/(\nu-2)}$ has the proper volatility.

Asymptotics of VaR_α and S_α: 1

- Given an independent sample, $\{L_i,\ 1 \leq i \leq n\}$, from the loss distribution, one can construct the empirical cdf.
- $F_L^{(n)}(x) = \frac{1}{n}(\text{number of } L_i \leq x)$.
- The Glivenko-Cantelli theorem guarantees that the empirical cdf will converge uniformly to the underlying cdf that generated the sample,

$$\sup_{-\infty < x < +\infty} |F_L^{(n)}(x) - F_L(x)| \to 0, \text{ as } n \to \infty.$$

Asymptotics of VaR_α and S_α: 2

- Applying the Glivenko-Cantelli idea to the inverse cdfs we find that the two risk measures will converge appropriately:

$$L_{(\lceil n\alpha \rceil)} \to F_L^{-1}(\alpha) = \text{VaR}_\alpha, \text{as } n \to \infty.$$

- Moreover,

$$S_\alpha^{(n)} \to \int_{\text{VaR}_\alpha}^{\infty} x \frac{dF_L(x)}{1-\alpha} \text{ as } n \to \infty.$$

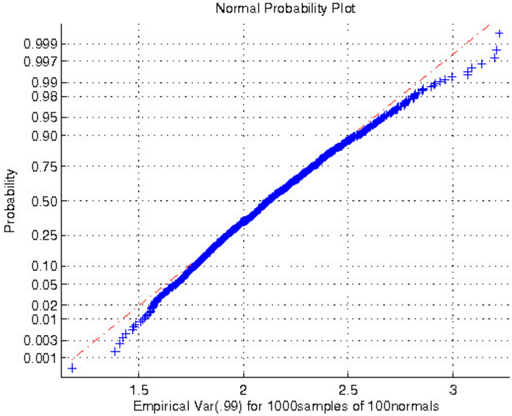
Asymptotics of VaR_α and S_α: 3

Unfortunately, the limiting variance of these estimates involves the underlying loss distribution. For example:

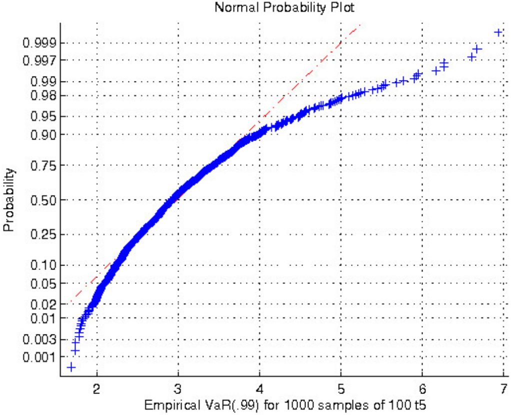
$$\begin{aligned}\sqrt{n}(L_{(\lceil n\alpha \rceil)} - \text{VaR}_\alpha) &\Rightarrow N(0, \frac{\alpha(1-\alpha)}{f_L^2(VaR_\alpha)}) \\ &= \frac{\sqrt{\alpha(1-\alpha)}}{f_L(\text{VaR}_\alpha)} N(0,1),\end{aligned}$$

where $f(VaR_\alpha)$ is the pdf of the loss distribution evaluated at the VaR_α value. More, later in the course. Additionally, the derivation of this result will be given in TA Session #1 on Friday, September 1, 11:00m.

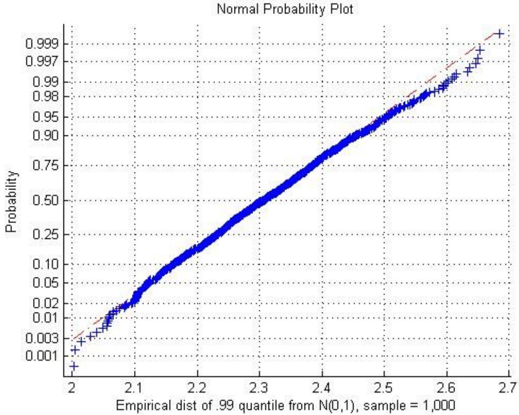
Example: Homework 1, Problem 2, VaR$_\alpha$



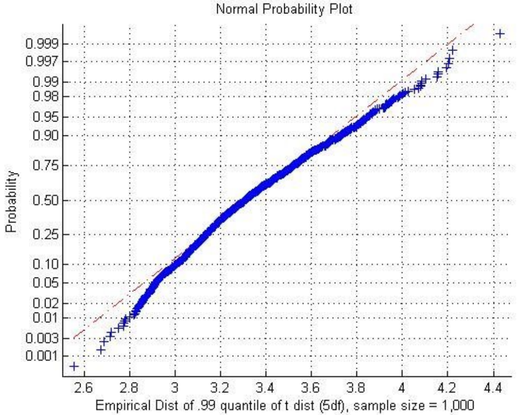
Example: Homework 1, Problem 2, VaR$_\alpha$



Example: Homework 1, Problem 2, VaR$_\alpha$



Example: Homework 1, Problem 2, VaR$_\alpha$



Coherent Risk Measures: 1

Artzner et. al. developed a set of axioms that a risk measure should follow. A risk measure that does so was called a coherent risk measure.

- Let L be the loss associated with a portfolio and let $\mathcal{R}(L)$ be the risk associated with that loss.
- Axiom 1: Monotonicity
  If L_1 and L_2 are two losses for which $P(L_1 \leq L_2) = 1$ (at least as much loss under all scenarios), then $\mathcal{R}(L_1) \leq \mathcal{R}(L_2)$.
- Axiom 2: Risk-free condition, or translation invariance
  Given any portfolio with associated loss L, suppose a riskless amount C is added to the portfolio, $-\infty < C < +\infty$. In this case $\mathcal{R}(\text{portfolio} \cup C) = \mathcal{R}(\text{portfolio}) - C$.

Coherent Risk Measures: 2

- Axiom 3: Positive homogeneity (scale change)
  Suppose $\lambda > 0$ for some constant lambda. Then
  $\mathcal{R}(\lambda L) = \lambda \mathcal{R}(L)$.
- Axiom 4: Subadditivity (diversification)
  Consider two portfolios with associated losses L_1 and L_2.
  Then $\mathcal{R}(L_1 + L_2) \leq \mathcal{R}(L_1) + \mathcal{R}(L_2)$.
- VaR does not satisfy Axiom 4 for every loss distribution; however, shortfall risk, S_α, is coherent.

Example of VaR "Incoherence": 1

VaR does not satisfy subadditivity. Here is one example:

- Consider a defaultable bond with current price of \$100, a 5% interest rate and a 1 year maturity. In one year there will be gain of \$5 with probability .98 or a loss of \$100 with probability of .02.
- If $.98 \leq \alpha \leq 1.00$, then $\text{VaR}_\alpha = 100$
  If $\alpha < .98$, then $\text{VaR}_\alpha = -5$

Example of VaR "Incoherence": 2

Consider n independent defaultable bonds with current price \$100, a 5% interest rate, and a 1 year maturity. The VaR value of each of these bonds is given by either 100 (if $\alpha \geq .98$) or -5 (if $\alpha < .98$).

- Consider a portfolio that is the union of n of these independent bonds. The sum of the VaR values, $\mathcal{R}(L_1) + \cdots \mathcal{R}(L_n)$ is given by 100n or -5n depending on whether $\alpha \geq .98$ or not.
- When the n bonds are combined into a single portfolio, then the loss, L, depends on how many defaults occur, X, a Binomial(n,.02) distributed random variable. $L = -5n + 105 * X$.
- For $(.98)^n < \alpha < .98$, $(X > 0)$, the subadditivity condition is reversed, i.e. the VaR value from n units of the same bond is smaller than the VaR value from holding n independent bonds with the same risk characteristics.

Example of VaR "Incoherence": 3

Here is a second example of VaR's lack of subadditivity. Consider two defaultable \$10M loans which will come due in one year. Each has probability of .9875 of not defaulting and paying off \$10,200,000, while each has probability .0125 of defaulting. Upon default, the bank will receive the recovery of the \$10M loan which is uniformly distributed over $[0, \$10M]$ (and the \$200,000 profit is lost). There is a dependency of default between the two loans, P(both default) = 0. Note the good diversification of the portfolio. Neither is very likely to default and at most one can default. Calculate the individual VaR and VaR of the portfolio.

Example of VaR "Incoherence": 4

Suppose we let $\alpha = .99$. Calculate the VaR of each individual loan and a portfolio consisting of both loans together.

- VaR will be -\$200,000 if $\alpha < .9875$ (no default). If there is a default (with probability .0125), then VaR will depend upon the recovery. The probability of default and recovery of 80% or greater (with probability .8) results in a loss of at most \$2M. The right tail of the cdf of *L* is linear connecting (0,.9875) and (10M,1.0). It reaches .99 on the y-axis where x = 2M. Consequently , the individual $VaR(.99) = \$2M$. The sum of the individual VaR values is \$4M.

Example of VaR "Incoherence": 5

- If the two loans are combined into one portfolio, then either neither defaults or one defaults but not both. VaR(.99) will be calculated in the latter case, $P(A \cup B) = P(A) + P(B) - P(A \cap B) = .0125 + .0125 - 0 = .0250$.
- If one fails, it pays off its recovery, while the other pays \$10.2M for a gain of \$200,000. If the recovery is 40% or less, then the total loss will be \$6M - \$200,000 = \$5.8M or more. The probability of a loss of \$5.8M or more is $.025 \times .4 = .01$.
- Consequently, VaR(.99) for the portfolio is \$5.8M.
- $\mathcal{R}(A) + \mathcal{R}(B) = \$2M + \$2M = \$4M \leq \$5.8 = \mathcal{R}(A \cup B)$, violating subadditivity.

Incoherence of VaR

VaR may fail the subadditivity condition when the underlying loss distribution is highly skewed or dominated by one or two large "elephants." In fact, VaR is subadditive for "elliptical distributions." For example,

- Supose there are two portfolios with associated losses, $L = (L_1, L_2)$ has $N_2((\mu_1, \mu_2), \Sigma)$ distribution.
- $\mathrm{VaR}_\alpha(L_i) = \mu_i + \sigma_i Z_\alpha, \;\; i = 1, 2.$
- $\mathrm{VaR}_\alpha(L_1 + L_2)) = \mu + \mu_2 + \sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2 Z_\alpha}.$
- Since for all $\rho \in [-1, 1]$: $\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2 \leq (\sigma_1 + \sigma_2)^2$
- Hence $\mathrm{VaR}_\alpha(L_1 + L_2) \leq \mathrm{VaR}_\alpha(L_1)$ + $\mathrm{VaR}_\alpha(L_1)$.

Approximations to VaR: 1

There are a few special cases where VaR can be calculated in closed form. Otherwise, VaR usually requires simulation or can be computed approximately. In the case of derivatives, one can approximate VaR using the Greeks.

- $L = V(\boldsymbol{S}, t) - V(\boldsymbol{S} + \Delta\boldsymbol{S}, t + \Delta)$ where $\boldsymbol{S}$ is the vector of current relevant risk factor values and $\Delta\boldsymbol{S}$ represents the change in risk factors.
- Expanding in a one-term multivariate Taylor series:

$$L \approx -\frac{\partial V}{\partial t}\Delta - \delta^T(\Delta\boldsymbol{S}),$$

  where $\delta = (\delta_1, \ldots, \delta_k)$ with $\delta_i = \frac{\partial V}{\partial S_i}$.
- This is known as the Delta-Normal Approximation
- This approximation is often part of the FRM exams.

Approximations to VaR: 2

Because the delta-normal approximation relies on both normal returns and near linear assets, one should include a second term in the Taylor expansion.

$$L \approx -\frac{\partial V}{\partial t}\Delta - \delta^T(\Delta \boldsymbol{S}) - \frac{1}{2}(\Delta \boldsymbol{S})^T \Gamma (\Delta \boldsymbol{S}),$$

where $\delta = (\delta_1, \dots, \delta_k)$ with $\delta_i = \frac{\partial V}{\partial S_i}$ and Γ is a $k \times k$ matrix whose elements are the gammas, $\Gamma_{ij} = \frac{\partial^2 V}{\partial S_i \partial S_j}$.

- This is known as the Delta-Gamma Approximation

Glasserman Delta-Gamma Figure

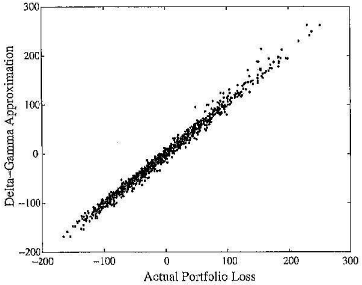
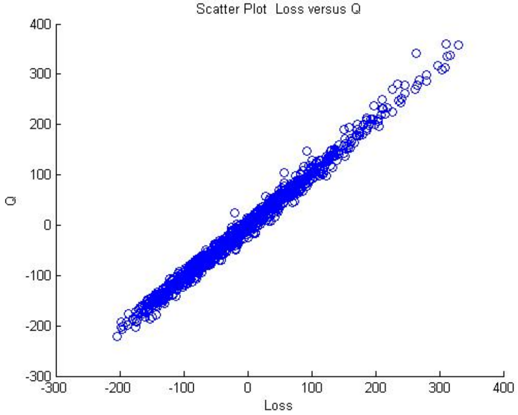
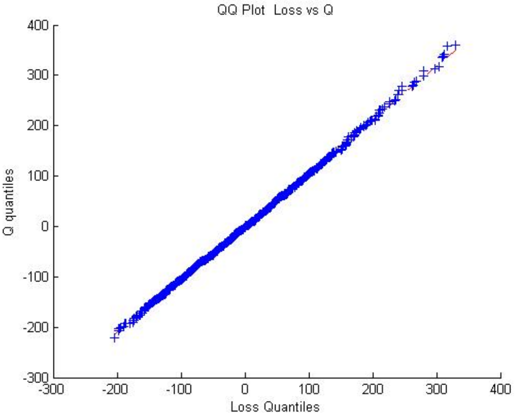


Fig. 9.1. Comparison of delta-gamma approximation and actual portfolio losses in 1000 randomly generated scenarios for a portfolio of 150 options.

Delta-Gamma Scatterplot



Delta-Gamma QQ-Plot



Backtesting: 1

- Risk measures like VaR and Expected Shortfall are not only management tools, they are also associated with regulation, especially in mandating the amount of capital a financial firm must have to cover its risk exposure.
- It is important that the risk system is working correctly. For example, if VaR is the risk measure, then there is a probability $(1-\alpha)$ that it will be violated, that is the actual loss experienced is larger than the VaR_α value.
- By looking over the sequence of VaR values and comparing them with the actual losses, one can perform a statistical test that the VaR calculations are being assessed properly.
- Backtesting the risk measure is achieved by comparing a series of pairs of values: the calculated VaR value with the actual loss value.

Backtesting: 2

- Since the loss is predicted assuming that a portfolio is not traded over $[t, t + \Delta]$, the actual loss can be calculated over a somewhat different portfolio. To make a correct comparison between predicted and observed, one must recalculate the value of the original portfolio rather than observing the value of the new portfolio.
- The Basel Accords mandated the use of VaR, and focused on its use in establishing capital requirements for financial institutions.
- To assess the adequacy of an institution's estimation of VaR, they developed a "stop light test."
- VaR should be calculated on a 1-day basis, the over a 1-year period, using $\alpha = .99$ the number of times that the actual loss exceeded the calculated VaR value (i.e. the number of violations or exceedances) was noted.
- For a Poisson approximation, $\lambda = 2.5$, so the number of violations has mean 2.5 and variance 2.5.

Backtesting: 3

The basic capital charge was assessed at 3 times the 10-day VaR. Depending upon the number of violations, 3 was increased to as much as 4. If VaR was consistently overestimated, Basel didn't care.

Zone	Number of Exceptions	Potential Increase in k
Green	0 to 4	0.00
Yellow	5	0.40
	6	.50
	7	.65
	8	.75
	9	.85
Red	$\geq$ 10	1.00

Backtesting: 4

- There are a variety of other testing approaches, but these mostly focus on the adequacy of the VaR, either being too low (resulting in exceptions) or too high (resulting in a firm having an unnecessarily large capital charge).
- These are two-sided tests, fundamentally based on the binomial distribution.
- The test by Kupiec (1995) which is a two-sided likelihood ratio test in which the null hypothesis is that there is a probability of α of the observed value being smaller than the VaR value and a probability of $1 - \alpha$ of it being larger. By counting the fractions of those above and of those below and comparing those with the value α yields a likelihood ratio test.

Backtesting: 5

- Basel is moving to expected shortfall; however, for quite some time it was thought that expected shortfall, while coherent, could not be backtested. Basel was going to allow expected shortfall to be used to calculate capital requirements but VaR for backtesting. This concern has been addressed and there are now a number of backtests of expected shortfall.
- Backtesting (VaR and Expected Shortfall) papers are on Canvas)

Extreme Value Theory

Extreme Value Theory: 1

There are two familiar and important theorems in probability theory that are of importance in computational finance: the law of large numbers, and the central limit theorem:

- Assume $L_1, L_2, \cdots, L_n$ are i.i.d. with some c.d.f. F, and assume that this distribution has a mean, m, and a standard deviation, $0 < \sigma < \infty$. Then, letting $S_n = \sum_{i=1}^{n} L_i$,
- Law of Large Numbers:

$$\frac{S_n}{n} = \bar{S}_n \to m.$$

- Central Limit Theorem :

$$\frac{\bar{S}_n - m}{\frac{\sigma}{\sqrt{n}}} \Rightarrow N(0,1).$$

Uses of Central Limit Theorem and Extremes

- The central limit theorem is very useful in giving good approximate answers to important questions.
- For example, how large a sample size is needed in estimating the price of a derivative security using Monte Carlo simulation if we find that each sample has a mean with the true price, *p*, and a standard deviation, η, and we want the estimated price to be within \$.05 of the true price?
- In risk management, we might want to estimate how large the largest drawdown will be in the next *n* trading days with some specified level of confidence.

Extreme Value Theory
Summary of Major Results

- Suppose we consider an i.i.d. sequence of random variables, $\{L_i,\ 1 \leq i \leq n\}$, each with cdf F. In risk management, these generally correspond to losses, and one is concerned with large values.
- In this situation, there are several results from probability theory that characterize the behavior of these large values are, therefore, of importance in risk management.
- These theorems fall into two categories: 1) the behavior of the largest sample value as $n \to \infty$ and 2) the behavior of the exceedances of a high value (like VaR): how often do they occur, what is their probability distribution, and what is their mean value? (= shortfall risk).

Extreme Value Theory: 2

There is a third important limit theorem that is relevant for risk management. It is associated with the maximum of a sample.

- Let $M_n = \max(L_1, \ldots, L_n)$. Suppose there is a sequence of constants, $\{a_n\}, \{b_n\}$ for which

$$\frac{M_n - a_n}{b_n} \Rightarrow H.$$

  where H is a non-degenerate cdf. Then H must take on one of three forms: the Fréchet distribution, the Weibull distribution, or the Gumbel distribution.

Extreme Value Theory: 3

- Fréchet (F has a polynomial tail)

$$H(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ e^{-x^{-\alpha}} & \text{if } x > 0 \end{cases}$$

- Weibull (F is bounded above, i.e. $F(x_0) = 1$ for some $x_0 < \infty$):

$$H(x) = \begin{cases} e^{-(-x^{-\alpha})} & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$$

Extreme Value Theory: 4

- Gumbel (F has an exponential tail)

$$H(x) = e^{-e^{-x}}, -\infty < x < \infty.$$

- These three distributions can be combined into a single representation called the Generalized Extreme Value Distribution (GEV):

$$H_\theta(x) = \begin{cases} e^{-(1+\theta x)^{-1/\theta}} & \text{if } \theta \neq 0 \\ e^{-e^{-x}} & \text{if } \theta = 0 \end{cases}$$

Extreme Value Theory: 5

Note, if we let $\theta = -1/n$, and substitute into the $\theta \neq 0$ expression for the GEV distribution:

$$\begin{aligned}
\lim_{n\to\infty}(H_\theta(x)) &= \exp(-\lim_{n\to\infty}(1+\theta x)^{-1/\theta}) \\
&= \exp(-\lim_{n\to\infty}(1-\frac{x}{n})^{n}) \\
&= e^{-e^{-x}} \\
&= H_0(x).
\end{aligned}$$

Extreme Value Theory: 6

Consider a simple example of a distribution, F, with an exponentially decreasing tail, namely one for which $\lim_{x\to\infty} e^{bx}(1 - F(x)) = c$. The exponential distribution is an example of a distribution having this property, since $1 - G(x) = e^{-\lambda x}$ and $e^{\lambda x}(1 - G(x)) = 1$ exactly, so $b = \lambda$ and $c = 1$.

$$\begin{aligned} P(M_n \le x) &= P(L_i \le x,\ 1 \le i \le n) \\ &= P^n(L_1 \le x) \\ &= (1 - e^{-\lambda x})^n. \end{aligned}$$

Extreme Value Theory: 7

Scaling and relocating, we find

$$
\begin{aligned}
P(\frac{M_n - a_n}{b_n} \le x) &= P(M_n \le a_n + b_n x) \\
&= (1 - e^{-\lambda(a_n + b_n x)})^n,
\end{aligned}
$$

and we need to choose a_n and b_n so this converges to a limiting cdf. Recall from elementary calculus:

$$
\lim_{n\to\infty} (1 + \frac{g(x)}{n})^n = e^{g(x)}.
$$

Extreme Value Theory: 8

$$P(\frac{M_n - a_n}{b_n} \leq x) = (1 - e^{-\lambda(a_n + b_n x)})^n.$$

First, we select $a_n = \frac{1}{\lambda}\log(n)$ which gives

$$P(\frac{M_n - a_n}{b_n} \leq x) = (1 - \frac{e^{-\lambda x b_n}}{n})^n.$$

If we next select $b_n = 1/\lambda$, the above expression becomes

$$(1 - \frac{e^{-x}}{n})^n \to e^{-e^{-x}}.$$

Extreme Value Theory: 9

Consequently,

$$\lambda M_n - \log(n) \to e^{-e^{-x}},$$

or

$$\lambda M_n \approx \log(n) + \text{Gumbel Noise},$$

and roughly $M_n \approx \frac{1}{\lambda}\log(n)$.

This gives the case of the exponential distribution. What if we consider the more general case where F has an exponential tail but may not be exactly the exponential distribution?

Extreme Value Theory: 10

We consider the case in which the underlying distribution, F, has the property that there are constants b and c for which

$$\lim_{x \to \infty} e^{bx}(1 - F(x)) = c.$$

$$\begin{aligned} P(\frac{M_n - a_n}{b_n} \leq x) &= P(M_n \leq a_n + b_n x) \\ &= F^n(a_n + b_n x) \\ &= (1 - (1 - F(a_n + b_n x)))^n \end{aligned}$$

Let $a_n = \frac{1}{b}\log(cn)$ and $b_n = \frac{1}{b}$. Thus
$\forall x,\ a_n + b_n x = \frac{1}{b}(\log(n) + \log(c) + x) \to \infty$ as $n \to \infty$.

Extreme Value Theory: 11

Note $a_n + b_n x = \frac{1}{b}(\log(n) + \log(c) + x) \to \infty$ as $n \to \infty$, and $e^{b(a_n+b_n x)} = nce^x$ It follows that:

$$\begin{aligned} P(\frac{M_n - a_n}{b_n} \le x) &= (1 - \frac{e^{b(a_n+b_n x)}}{e^{b(a_n+b_n x)}}(1 - F(a_n + b_n x)))^n \\ &\approx (1 - \frac{c}{cne^x})^n \\ &= (1 - \frac{e^{-x}}{n})^n \\ &\to e^{-e^{-x}} \end{aligned}$$

which is the Gumbel cdf.

Extreme Value Theory: 12

In addition to the extreme value distributions, individually or in their consolidated form, there is another distribution that is relevant to risk management, the Generalized Pareto Distribution (GPD):

$$G_{\theta,\beta} = \begin{cases} 1 - (1 + \theta x/\beta)^{-1/\theta} & \text{if } \theta \neq 0 \text{ and } x \geq 0 \\ 1 - e^{-x/\beta} & \text{if } \theta = 0 \text{ and } 0 \leq x \leq -\beta/\theta \end{cases}$$

where $\beta > 0$ is a scale parameter, and $E(X) = \beta/(1 - \theta)$, $\theta < 1$.