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

Motivation

- Consider a Loss with a distribution $L \sim \mathcal{N}(0, 1)$.
- (Q1) What is the maximum loss? (∞)
- (Q2) What is the maximum loss with 95% confidence? (1.645)
- Q2 is the 95% Value-at-Risk (VaR) for the loss distribution.

VaR and Regulatory Capital

- Regulators have traditionally used VaR to calculate the capital banks are required to maintain.
- Market-risk capital is typically based on a 10-day VaR estimated at a 99% confidence level.
- Credit risk and operational risk capital are generally based on a one-year 99.9% VaR.

Advantages of VaR

- Captures an important aspect of risk in a single number.
- Easy to understand and communicate.
- Answers the simple question: "How bad can things get?"

Value-at-Risk (VaR): Definition

- **Value-at-Risk (VaR)** at confidence level $\alpha \in (0,1)$:

$$\begin{aligned}\mathrm{VaR}_\alpha(L) &= \inf\{x \in \mathbb{R} : P(L \leq x) \geq \alpha\} \\ &= \inf\{x : F_L(x) \geq \alpha\} \\ &= \inf\{x : P(L > x) \leq 1 - \alpha\}\end{aligned}$$

- $\mathrm{VaR}_\alpha(L) = F_L^{-1}(\alpha)$ is the α-quantile of the loss distribution, or generalized inverse of the CDF F_L.
- Interpretation: the maximum loss not exceeded with probability α.
- Typical values for α: 0.95, 0.99.

Example: VaR Calculation

- The gain from a portfolio during six months is normally distributed with mean 2 million and standard deviation 10 million.

$$G \sim \mathcal{N}(2, 10^2), L = -G$$

Example: VaR Calculation

- The gain from a portfolio during six months is normally distributed with mean 2 million and standard deviation 10 million.

$$G \sim \mathcal{N}(2, 10^2), L = -G$$

- The 1% point of the distribution of gains is $2 - 2.33 \times 10 = -21.3$ million.

$$\mathbb{P}(G \leq -21.3) = 1\% \text{ or } F_G^{-1}(0.01) = (-21.3)$$

- The VaR for the portfolio with a six-month time horizon and a 99% confidence level is 21.3 million.

$$\mathrm{VaR}_{0.99}(L) = 21.3 \text{ million}$$

Example: VaR Calculation

- The gain from a portfolio during one year is uniformly distributed between -50 million and 50 million.

$$G \sim \mathcal{U}(-50, 50), L = -G$$

Example: VaR Calculation

- The gain from a portfolio during one year is uniformly distributed between -50 million and 50 million.

$$G \sim \mathcal{U}(-50, 50), L = -G$$

- The 1% point of the distribution of gains is $-50 + 0.01 \times 100 = -49$ million.

$$\mathbb{P}(G \leq -49) = 1\% \text{ or } F_G^{-1}(0.01) = -49$$

- The VaR for the portfolio with a one-year time horizon and a 99% confidence level is 49 million.

$$\mathrm{VaR}_{0.99}(L) = 49 \text{ million}$$

Example: VaR Calculation

- A one-year project has:
  - A 98% chance of leading to a gain of 2 million,
  - A 1.5% chance of a loss of 4 million,
  - A 0.5% chance of a loss of 10 million.

Example: VaR Calculation

- A one-year project has:
  - A 98% chance of leading to a gain of 2 million,
  - A 1.5% chance of a loss of 4 million,
  - A 0.5% chance of a loss of 10 million.
- The VaR with a 99% confidence level is 4 million.
- What if the confidence level is 99.9%?
- What if the confidence level is 99.5%?

Example: VaR Calculation

- Suppose L takes values $\{0, 1, 2, 3\}$ with probabilities $\{0.5, 0.3, 0.15, 0.05\}$. Then:
  - At a confidence level of $\alpha = 0.95$, $\mathrm{VaR}_{0.95}(L) = 2$.
  - At a confidence level of $\alpha = 0.8$, $\mathrm{VaR}_{0.8}(L) = 1$.

VaR properties: Monotonicity

- If $L_1 \leq L_2$, then $\mathrm{VaR}_\alpha(L_1) \leq \mathrm{VaR}_\alpha(L_2)$.
- (proof)
  - Set $x_i = \mathrm{VaR}_\alpha(L_i) = \inf\{y : \mathbb{P}(L_i \leq y) \geq \alpha\}$.
  - $\mathbb{P}(L_i \leq y) \geq \alpha$ iff $y \geq \mathrm{VaR}_\alpha(L_i)$.
  - $L_1 \leq L_2$ implies $\{L_1 \leq x_2\} \supseteq \{L_2 \leq x_2\}$.
  - $\mathbb{P}(L_2 \leq x_2) \geq \alpha$ implies $\mathbb{P}(L_1 \leq x_2) \geq \alpha$.
  - Thus, $\mathrm{VaR}_\alpha(L_1) \leq x_2$.
- Interpretation: Monotonicity ensures that if one loss amount is always less than or equal to another, the VaR of the first will not exceed the VaR of the second.

VaR properties: Translation Invariance

- For any constant c, $\mathrm{VaR}_\alpha(L + c) = \mathrm{VaR}_\alpha(L) + c$.
- (proof)
  - Let $x = \mathrm{VaR}_\alpha(L)$, so $\mathbb{P}(L \leq x) \geq \alpha$.
  - For $L + c$, $\mathbb{P}(L + c \leq x + c) = \mathbb{P}(L \leq x) \geq \alpha$.
  - Thus, $\mathrm{VaR}_\alpha(L + c) \geq x + c = \mathrm{VaR}_\alpha(L) + c$.
  - Take $\hat{L} = L + c$ and $\hat{c} = -c$, we have $\mathrm{VaR}_\alpha(\hat{L} + \hat{c}) \geq \mathrm{VaR}_\alpha(\hat{L}) + \hat{c}$. This implies that $\mathrm{VaR}_\alpha(L) \geq \mathrm{VaR}_\alpha(L + c) - c$.
- Interpretation: Adding a constant to the loss shifts the VaR by the same constant.

VaR properties: Positive Homogeneity

- For any $\lambda > 0$, $\mathrm{VaR}_\alpha(\lambda L) = \lambda \mathrm{VaR}_\alpha(L)$.
- (proof)
  - Let $x = \mathrm{VaR}_\alpha(L)$, so $\mathbb{P}(L \leq x) \geq \alpha$.
  - For λL, $\mathbb{P}(\lambda L \leq \lambda x) = \mathbb{P}(L \leq x) \geq \alpha$.
  - Thus, $\mathrm{VaR}_\alpha(\lambda L) \geq \lambda \mathrm{VaR}_\alpha(L)$.
  - Conversely, let $\hat{L} = \lambda L$ and $\hat{\lambda} = \lambda^{-1}$, we have $\mathrm{VaR}_\alpha(\hat{\lambda}\hat{L}) \geq \hat{\lambda}\mathrm{VaR}_\alpha(\hat{L})$. This implies that $\mathrm{VaR}_\alpha(L) \geq \lambda^{-1}\mathrm{VaR}_\alpha(\lambda L)$.
- Interpretation: Scaling the loss by a positive factor scales the VaR by the same factor.

VaR for normal distributions

- If L has a continuous and strictly increasing distribution function F_L, then the Value-at-Risk can be expressed as:

$$\mathrm{VaR}_\alpha(L) = F_L^{-1}(\alpha)$$

- If $L \sim N(\mu, \sigma^2)$, then:

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

where Φ^{-1} is the standard normal quantile function.

VaR for Student's t-distribution

- For $L \sim t(\nu, \mu, \sigma^2)$ (location-scale t-distribution):

$$\mathrm{VaR}_\alpha(L) = \mu + \sigma \cdot t_\nu^{-1}(\alpha)$$

  where $t_\nu^{-1}(\alpha)$ is the α-quantile of the standard t-distribution with ν degrees of freedom.

Expected Shortfall (ES): Definition

- **Expected Shortfall (ES)** at confidence level α:

$$\mathrm{ES}_\alpha(L) = \frac{1}{1-\alpha} \int_\alpha^1 F_L^{-1}(u)\, du$$

Expected Shortfall (ES): Interpretation

- For $L \in \mathbb{L}^1$ with continuous F_L, we have:

$$\mathrm{ES}_\alpha(L) = \mathbb{E}[L \mid L > \mathrm{VaR}_\alpha(L)]$$

- Interpretation: the average loss in the worst α fraction of cases.
- ES is also known as Conditional Value-at-Risk (CVaR).
- It is not always true, especially for discrete distributions.

Proof: ES as Conditional Expectation

- We show that for continuous F_L:

$$\mathrm{ES}_\alpha(L) = \mathbb{E}[L \mid L > \mathrm{VaR}_\alpha(L)]$$

- Let $q = \mathrm{VaR}_\alpha(L) = F_L^{-1}(\alpha)$. Then:

$$\begin{aligned}
\mathbb{E}[L \mid L > q] &= \frac{\mathbb{E}[L \cdot \mathbf{1}_{\{L>q\}}]}{P(L > q)} \\
&= \frac{1}{1-\alpha} \int_q^\infty x dF_L(x) \\
&= \frac{1}{1-\alpha} \int_\alpha^1 F_L^{-1}(u) du
\end{aligned}$$

- The last equality uses the substitution $u = F_L(x)$.

Example: ES Calculation

- Let L take values $\{0, 1, 2, 3, 4\}$ with probabilities $\{0.4, 0.3, 0.2, 0.08, 0.02\}$.
- For $\alpha = 0.9$:
  - $P(L \leq 2) = 0.9$, so $\mathrm{VaR}_{0.9}(L) = 2$
  - Values exceeding VaR: $\{3, 4\}$ with probabilities $\{0.08, 0.02\}$
  - $\mathrm{ES}_{0.9}(L) = \frac{1}{0.1}[3 \times 0.08 + 4 \times 0.02] = \frac{0.32}{0.1} = 3.2$
- For $\alpha = 0.95$, we have:
  - $\mathrm{VaR}_{0.95}(L) = 3$
  - $\mathrm{ES}_{0.95}(L) = \frac{1}{0.05}[4 \times 0.02 + 3 \times (0.05 - 0.2)] = \frac{0.17}{0.05} = 3.4$

ES for Normal Distribution

- For $L \sim N(\mu, \sigma^2)$ and confidence level α:

$$\mathrm{ES}_\alpha(L) = \mu + \sigma \cdot \frac{\varphi(\Phi^{-1}(\alpha))}{1-\alpha}$$

where φ is the standard normal PDF and Φ^{-1} is the standard normal quantile function.

Proof: ES for Normal Distribution

- For $L \sim N(\mu, \sigma^2)$, we derive $\mathrm{ES}_\alpha(L)$.
- Let $Z = \frac{L-\mu}{\sigma} \sim N(0,1)$. Then $L = \mu + \sigma Z$.
- We have:

$$\begin{aligned}\mathrm{ES}_\alpha(L) &= \mathbb{E}[L \mid L > \mathrm{VaR}_\alpha(L)] \\ &= \mathbb{E}[\mu + \sigma Z \mid Z > \Phi^{-1}(\alpha)] \\ &= \mu + \sigma \mathbb{E}[Z \mid Z > \Phi^{-1}(\alpha)]\end{aligned}$$

- For $Z \sim N(0,1)$ and $z_\alpha = \Phi^{-1}(\alpha)$:

$$\begin{aligned}\mathbb{E}[Z \mid Z > z_\alpha] &= \frac{1}{1-\alpha} \int_{z_\alpha}^{\infty} z\varphi(z)dz \\ &= \frac{\varphi(z_\alpha)}{1-\alpha}\end{aligned}$$

- Therefore: $\mathrm{ES}_\alpha(L) = \mu + \sigma \cdot \frac{\varphi(\Phi^{-1}(\alpha))}{1-\alpha}$

Example: VaR and ES for Normal Losses

- $L \sim N(0,1)$, $\alpha = 0.99$
- $\mathrm{VaR}_{0.99}(L) = \Phi^{-1}(0.99) \approx 2.33$
- $\mathrm{ES}_{0.99}(L) = \frac{\varphi(\Phi^{-1}(0.99))}{1-0.99} \approx 2.67$
- φ is the standard normal density.

VaR for Student's t-distribution

- For $L \sim t(\nu, \mu, \sigma^2)$ (location-scale t-distribution):

$$\mathrm{VaR}_\alpha(L) = \mu + \sigma \cdot t_\nu^{-1}(\alpha)$$

  where $t_\nu^{-1}(\alpha)$ is the α-quantile of the standard t-distribution with ν degrees of freedom.

Coherent Risk Measures: Intuition

- **Monotonicity:** If one portfolio always produces a worse outcome than another, its risk measure should be greater.
- **Translation Invariance:** Adding an amount of cash K to a portfolio should reduce its risk measure by K.
- **Positive Homogeneity:** Scaling the size of a portfolio by λ should scale the risk measure by λ.
- **Subadditivity:** The risk measure for two portfolios after they are merged should be no greater than the sum of their individual risk measures.

Coherent Risk Measures

- A risk measure ρ is **coherent** if it satisfies:
  1. **Monotonicity:** If $L_1 \leq L_2$ a.s., then $\rho(L_1) \leq \rho(L_2)$.
  2. **Subadditivity:** $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$.
  3. **Translation invariance:** For $a \in \mathbb{R}$, $\rho(L + a) = \rho(L) + a$.
  4. **Positive homogeneity:** For $\lambda \geq 0$, $\rho(\lambda L) = \lambda\rho(L)$.
- VaR is not coherent due to lack of subadditivity.
- ES is always subadditive; ES is coherent.

Generalized quantile function

- For a random variable X with CDF F_X, the generalized quantile function is defined as VaR:

$$F_X^{-1}(\alpha) = \mathrm{VaR}_\alpha(X)$$

- Properties:
    - $F^{-1}(\alpha)$ is left-continuous and non-decreasing
    - If F is continuous and strictly increasing, then $F^{-1}(F(\alpha)) = \alpha$
    - $F(F^{-1}(\alpha)) \geq \alpha$ always holds

Properties of the quantile function

Let X be a random variable with distribution function F_X.

- (Quantile transformation) If $U \sim U(0,1)$, then $F_X^{-1}(U)$ has distribution F_X.
- (Probability transformation) If F_X is continuous, then $F_X(X) \sim U(0,1)$.

Comonotonicity

- $X_1, \ldots, X_n$ are comonotonic if $\exists\ Z$ and increasing functions f_i s.t. $X_i = f_i(Z)$.
- Represents the strongest possible positive dependence.
- No diversification benefit: $\mathrm{VaR}_\alpha(\sum X_i) = \sum \mathrm{VaR}_\alpha(X_i)$.
- Important for worst-case risk aggregation.

Proof: No Diversification for Comonotonic Variables

Theorem

If $X_1, \ldots, X_n$ are comonotonic, then $\mathrm{VaR}_\alpha(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathrm{VaR}_\alpha(X_i)$.

Proof.

Since $X_i = f_i(Z)$ with f_i increasing:

$$P\left(\sum_{i=1}^n X_i \leq x\right) = P\left(\sum_{i=1}^n f_i(Z) \leq x\right) = P\left(Z \leq \left(\sum_{i=1}^n f_i\right)^{-1}(x)\right)$$

For the α-quantile:

$$\begin{aligned}\mathrm{VaR}_\alpha\left(\sum_{i=1}^n X_i\right) &= \left(\sum_{i=1}^n f_i\right)\left(F_Z^{-1}(\alpha)\right) \\ &= \sum_{i=1}^n f_i\left(F_Z^{-1}(\alpha)\right) = \sum_{i=1}^n \mathrm{VaR}_\alpha(X_i)\end{aligned}$$

Example: Positive Dependence but Not Comonotonic

The loss variables X and Y are given by:

$$X = Z + \rho\epsilon, \quad Y = Z + 2\rho\epsilon$$

where $Z, \epsilon \sim N(0,1)$ independent, and some constant ρ.

1. Is X, Y positively dependent?
2. For $\alpha = 0.99$ and $\rho = 1/2$, compute $\mathrm{VaR}_\alpha(X+Y)$ and compare with $\mathrm{VaR}_\alpha(X) + \mathrm{VaR}_\alpha(Y)$.
3. Are X, Y comonotonic?