Applications: risk measure estimation, backtesting, and extensions.

Camilo A. García Trillos

Market risk and portfolio theory

Risk measures in the market-risk management

Market risk management requires the frequent estimation of the incurred risks in all market-related positions of a financial company.

Some examples of application:

- Calculation of capital
- Margin calculation
- Reinsurance pricing

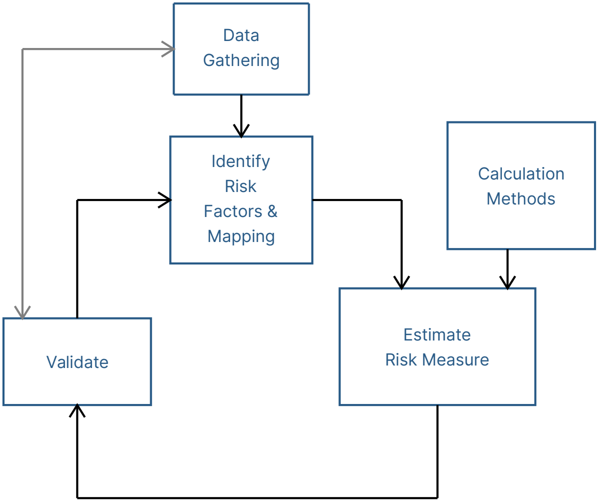
Risk measures in time

Risk measures are calculated in practice at a given date for possibly several periods, **using the information known at the date of calculation**.

This means:

- A filtration is involved (technically we need to calculate *conditional risk measures*)
- Multi-period risk measures or risk values for different periods need to be aggregated.

Risk measure estimation



- Sometimes we also *aggregate* estimations

Risk mapping

1. Transform quantities to simplify expressions (observable factors);
2. To account for incomplete data ⚠(observable and unobservable factors)

Mapping of risks: Examples

Set $\Delta V_{k+1} = V_{k+1} - V_K$ the P&L of an investment.

- Example 1: Stock portfolio Take increments in log-prices of quoted stocks as risk factors, that is $F^i_{k+1} = \log(S^i_{k+1}) - \log(S^i_k)$. Then

$$\Delta V_{k+1} = \sum_{i=1}^{n} \pi_i V_k (e^{F^i_{k+1}} - 1)$$

Mapping of risks: Examples

Set $\Delta V_{k+1} = V_{k+1} - V_K$ the P&L of an investment.

- Example 1: Stock portfolio Take increments in log-prices of quoted stocks as risk factors, that is $F^i_{k+1} = \log(S^i_{k+1}) - \log(S^i_k)$. Then

$$\Delta V_{k+1} = \sum_{i=1}^{n} \pi_i V_k (e^{F^i_{k+1}} - 1)$$

- Example 2: Bond portfolio Take as risk factor the continuously compounded yield curve $y(s, T) := -(T - s)^{-1} \log P(s, T)$

$$\begin{aligned}\Delta V_{k+1} &= \sum_{i=1}^{n} V_k \pi_i (P(t_{k+1}, T_i) - P(t_k, T_i)) \\ &= \sum_{i=1}^{n} \pi_i V_k \{e^{-(T - t_{k+1})[y(t_{k+1}, T) - y(t_k, T)] + y(t_k, T)\Delta} - 1\}\end{aligned}$$

Estimation

Depending on speed, accuracy and data constraints, we have as choices

1. Sensitivity method
2. Monte Carlo simulations
3. Historical simulations

Sensitivity method - Gaussian case

Suppose that,

$$\Delta V_{k+1} = \sigma_k \xi_{k+1} + m_k$$

where $(\xi_k)_{k\in\mathbb{N}}$ is stationary and Gaussian with $\xi_k \sim \mathcal{N}(0,1)$.

Then we can use known properties of normal random variables, to deduce, for example:

- $\mathrm{V@R}_k^\alpha(\Delta V_{k+1}) = \sigma_k \Phi^{-1}(\alpha) - m_k$,
- $\mathrm{ES}_k^\alpha(\Delta V_{k+1}) = \sigma_k \frac{\varphi(\Phi^{-1}(\alpha))}{1-\alpha} - m_k$.

Thus both risk measures would be available as soon as we estimate m_k, σ_k.

Continuity of $\mathrm{V@R}$ and ES

We can show that if a sequence $\{X_n\} \xrightarrow{d} X$. Recall also that convergence in L^2 and almost sure imply convergence in law. We get,

$$\mathrm{V@R}^{\alpha}(X_n) \to \mathrm{V@R}^{\alpha}(X).$$

As a direct consequence, we get

$$\mathrm{ES}^{\alpha}(X_n) \to \mathrm{ES}^{\alpha}(X).$$

Δ approximation

Take a factor model with stationary Gaussian risk increments, i.e.

$$V_{k+1} = f(\xi^1_{k+1}, \ldots, \xi^\ell_{k+1}) = f(\boldsymbol{\xi}_{k+1})$$

with $\Delta\boldsymbol{\xi}_{k+1} := \boldsymbol{\xi}_{k+1} - \boldsymbol{\xi}_k \sim \mathcal{N}(\boldsymbol{m}, \bar{\Sigma})$. If h is differentiable, then

$$\Delta V_{k+1} = (\Delta\xi^1_{k+1}, \ldots, \Delta\xi^\ell_{k+1}) \cdot \nabla_\xi f(\xi^1_k, \ldots, \xi^\ell_k) + O(|\Delta\boldsymbol{\xi}_{k+1}|^2)$$

We can then use the closed-form for Gaussian in order to **approximate** $\mathrm{V@R}$ and ES

Δ approximation (cont.)

Example 1: Stock Portfolio With $\boldsymbol{F}_k$ the vector of log-returns we have using linearisation that

$$V_{k+1} = V_k \sum_{i=1}^{n} \pi_i \exp(F_{k+1}^i) \Rightarrow \Delta V_{k+1} = V_k \sum_{i=1}^{n} \pi_i \Delta F_{k+1}^i + O(\|\Delta F_{k+1}\|^2)$$

If the data are stationary and Gaussian ($\Delta \boldsymbol{F}_k \sim \mathcal{N}(\boldsymbol{m}, \bar{\Sigma})$), then

$$\mathrm{V@R}_k^\alpha(\Delta V_{k+1}) \approx V_k(\boldsymbol{\pi}^\top \bar{\Sigma} \boldsymbol{\pi}) \Phi^{-1}(\alpha) - V_k(\boldsymbol{\pi} \cdot \boldsymbol{m})$$

(similarly for ES).

Δ approximation (cont.)

Example 2: One stock, GARCH model

Suppose that increments in log returns follow a GARCH(1,1)

$$
\begin{aligned}
\Delta F_{k+1} &= a_0 + a_1 \Delta F_k + \sigma_k \xi_{k+1} \\
\sigma_k^2 &= b_0 + b_1 \sigma_{k-1}^2 + b_2 \xi_k^2
\end{aligned}
$$

Where $(\xi_k)_{k \in \mathbb{N}}$ are i.i.d and standard Gaussian.
As long as we can estimate the parameters, we can define

$$m_k := a_0 + a_1 \Delta F_k$$

so that

$$\mathrm{V@R}_k^{\alpha}(\Delta V_{k+1}) \approx \sigma_k V_k \Phi^{-1}(\alpha) - m_k V_k,$$

The estimation step can be done using statistical techniques (for example maximum likelihood) from historical data.

Advantages of Δ approximation

Advantages:

- Fast calculation: Only requires the “sensitivities” (derivatives) of the function, and estimating mean and variance of the associated process.
- Easy to understand the relative importance of each risk factor.
- Not limited to Gaussian : any distribution for which we have closed form expressions for risk measures (for example log-normal, t-distribution, normal-Poisson mixture, ...).
- Can also be generalised to second order $\gamma - \Delta$ approximation (but with stronger assumptions).

Disadvantages of Δ approximation

Disadvantages:

- Requires differentiability of the risk mapping.
- Strong assumptions: Data should be able to show the assumed behaviour.
- Not very accurate: The quadratic term error that we are neglecting can be very large. In practice only valid for small perturbations (for example small-time periods).

Risk measure estimation: Monte Carlo

We can avoid making linearity/normality assumptions if we are willing to loose closed form expressions.

In this case:

- Choose parametric model for risk-factors $F^1 \ldots, F^n$.
- Calibrate distributions to market information or stress-scenario considerations
- Generate a large number M of samples $(F_k^{i,(1)}, \ldots, F_k^{i,(M)})$.
- Calculate $\Delta V_k^{(j)} = f(t_{k+1}, F_{k+1}^{1,(j)} \ldots, F_{k+1}^{n,(j)}) - f(t_k, F_k^{1,(j)} \ldots, F_k^{n,(j)})$
- Obtain risk measure for $\Delta V^{(j)}$ from empirical distribution.

Risk measure estimation: Monte Carlo (cont.)

Recall the empirical estimation of the risk measures: assume samples are ordered, that is $\Delta V_k^{(i)} \leqslant \Delta V_k^{(j)}$ if $i \leqslant j$:

Example The following are estimators of $\mathrm{V@R}^{\alpha}$:

$$\mathrm{V@R}^{\alpha}_{k,+} := -\Delta V_k^{(\lfloor (1-\alpha) M \rfloor)}; \qquad \mathrm{V@R}^{\alpha}_{k,-} := -\Delta V_k^{(\lceil (1-\alpha) M \rceil)};$$

$$\begin{aligned} \mathrm{V@R}^{\alpha}_{k,mid} := & -\left(M\alpha - \lfloor \alpha M \rfloor\right) \mathrm{V@R}^{\alpha}_{k,+} \\ & - \left(\lceil \alpha M \rceil - M\alpha\right) \mathrm{V@R}^{\alpha}_{k,-} \end{aligned}$$

while the following is an estimator for ES^{α}

$$\mathrm{ES}^{\alpha}_{k,+} = \frac{-1}{\lfloor (1-\alpha) M \rfloor} \sum_{i=1}^{\lfloor (1-\alpha) M \rfloor} \Delta V_k^{(i)}$$

Monte Carlo is a trade-off

Advantages:

- Can be applied to very general models
- Allows to control the balance between accuracy and calculation time
- Allows the introduction of stressed scenarios not available on the data

Disadvantages:

- Slow calculation, specially when the risk mappings are complex
- Requires a model
- A priori, it is difficult to establish importance of each risk measure / allocate risks.

Historical simulation

This is essentially a version of the Monte Carlo simulation, where we use the **empirical distribution** obtained from the data.

Key point: Estimation of empirical distribution for the risk factors

As we have seen, this implies that the risk factors should be stationary. We compose mappings to achieve this.

Historical simulation (cont.)

Example: Assume that

$$F_{k+1} = F_k + \xi_{k+1},$$

for stationary $(\xi_k)_{k\in\mathbb{N}}$ with unknown distribution.
Then, we define for $i = 1, \ldots, H$ (H is history size)

$$\Delta V_k^{(i)} := f(F_k^1 + (F_{k-i+1}^1 - F_{k-i}^1), \ldots, F_k^1 + (F_{k-i+1}^n - F_{k-i}^n)) - V_k$$

and apply the estimators as in the Monte Carlo case

Historical Simulation is a trade-off

Advantages:

- No a priori assumption on the distribution
- Data driven
- Straightforward connection with backtesting (...)

Historical Simulation is a trade-off

Disadvantages:

- Requires excellent data quality (less robust to noise in data)
- Less general models than Monte Carlo (due to stationarity requirements).
- Slowest calculation, specially when the risk mappings are complex
- A priori, it is difficult to establish importance of each risk measure / allocate risks.
- Total reliance on past events.

To summarise...

Closed form: Fastest. Intuitive. Extremely strong assumptions. Inaccurate. Backward looking.

Monte Carlo: Slow. Might be difficult to interpret some results. Requires modelling. Strong assumptions. Allows for forward looking.

Historical: Slowest. Results can be interpreted but not intuitive. Requires stationarity. Backward looking.

Validation

Regardless of the method we use to calculate our risk measures, we only obtain **estimators** subject to both numerical and model errors.

We need to validate these results.

Definition (Backtesting)

The use of some *statistically meaningful tests* and *past observed realisations* to evaluate the goodness of our risk estimation.

Reminder of hypothesis testing

Example: Claim (H_0): The random variable Z is normal with mean 1 and variance 4.

We observe $N = 10^6$ (one million) i.i.d samples of Z, with sample average 1.05.

Question: Is my initial claim reasonable given the observed data?

Reminder of hypothesis testing

Example: Claim (H_0): The random variable Z is normal with mean 1 and variance 4.

We observe $N = 10^6$ (one million) i.i.d samples of Z, with sample average 1.05.

Question: Is my initial claim reasonable given the observed data?

Refined question: Under the assumption, claim is **unreasonable** if it is *very unlikely* to observe an even greater distance from the mean.

Reminder of hypothesis testing

Solution: By the central limit theorem, we have

$$\hat{Z}_1 := \sqrt{N}\frac{(\hat{Z} - \mathbb{E}[Z])}{\sqrt{\mathrm{var[Z]}}} \xrightarrow{d} \mathcal{N}(0, 1)$$

where $\hat{Z}$ is the sample mean r.v. In our case,

$$P[|\hat{Z}| > \hat{z}] = \mathbb{P}[|\hat{Z}_1| > a]$$

where

$$a = \frac{\sqrt{10^6} \times (1.05 - 1)}{\sqrt{4}} = 25.$$

Hence, $P[|\hat{Z}| > \hat{z}] \approx 6.113 \times 10^{-138}$!!!

Thus, either our claim is false or we are observing an extremely unlikely event. We *reject the null hypothesis*.

Reminder of hypothesis testing

Let's revisit the question, this time assuming we observe $N = 100$ i.i.d. samples of Z with sample average 1.05 as before. We get

$$a = \frac{\sqrt{10^2} \times (1.05 - 1)}{\sqrt{4}} = 0.25$$

and thus, $P[|\hat{Z}| > \hat{z}] \approx 0.8026$.

This does not seem as an unlikely event any more.

Question: Can we conclude that the null hypothesis holds?

Reminder of hypothesis testing

		Decision	
		Retain null	Reject null
Truth in population	True	Correct: $(1 - \alpha)$	Type I error: α
	False	Type II error	Correct

- Unless an alternative is considered, we focus on obtaining evidence to reject the null assumption (small type I error), but not on obtaining evidence to support it ⚠

Reminder of hypothesis testing

		Decision	
		Retain null	Reject null
Truth in population	True	Correct: $(1-\alpha)$	Type I error: α
	False	Type II error: β	Correct (power) : $(1-\beta)$

- Unless an alternative is considered, we focus on obtaining evidence to reject the null assumption (small type I error), but not on obtaining evidence to support it ⚠
- If an alternative assumption is available, we can also control the type II error by choosing the number of samples and statistics.

Reminder of hypothesis testing

To summarise: To perform the test

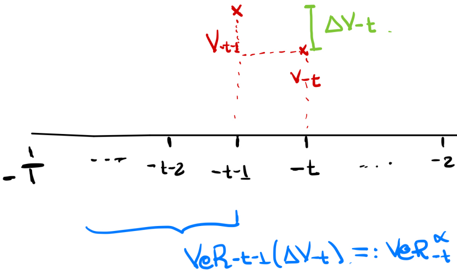
1. State the hypotheses (null hypothesis, H_0)
2. Set the criteria for decision:
   - Estimator
   - Reference probability for rejection α
   - If alternative assumption available fix also β
   - Type of test (two-tailed, left-tailed or right-tailed)
3. Compute the test statistic and its p-value
4. Make a decision: if p-value is smaller than reference, reject the null hypothesis.

Some properties that can be backtested

Some properties that are desired from our estimation:

- Coverage or adequacy property
- No clustering
- Comparative effectiveness

Framework for backtesting of $\mathrm{V@R}$, ES



- ΔV_{-t}: P&L in $[-t-1, -t]$
- $\mathrm{V@R}^{\alpha}_{-t}$: estimated value at risk at level α. (similarly $\mathrm{ES}^{\alpha}_{-t}$)
- $I_{-t} := \mathbb{1}_{\{-\Delta V_{-t} \geqslant \mathrm{V@R}^{\alpha}_{-t}\}}$: observed excess over predicted $\mathrm{V@R}$.

Backtesting $\mathrm{V@R}$ (unconditional)

Null hypothesis

$$H_0 : \text{The sequence } \{I_{-t}\}_{t=1,\dots,T} \text{ is i.i.d. and } \mathbb{P}[I_{-t} = 1] = (1-\alpha),$$

(no clustering on $\mathrm{V@R}$ breaches and $\mathrm{V@R}$ is well calculated).

Remark: For now, we *hold* the assumption of the independence of samples, and only test for coverage of $\mathrm{V@R}$. We later come back to the independence question.

Under H_0:

- $Z_{\mathrm{V@R},1} = \sum_{t=1}^{T} \hat{I}_{-t}$; follows the binomial distribution with T steps and $(1-\alpha)$ probability.

Backtesting $\mathrm{V@R}$ (unconditional)

Likewise, under H_0:

- $Z_{\mathrm{V@R},2} = \frac{\sqrt{T}}{\sqrt{\alpha(1-\alpha)}}\left(\frac{Z_{\mathrm{V@R},1}}{T} - (1-\alpha)\right)$; is asymptotically standard Gaussian.
- $Z_{\mathrm{V@R},3} = 2\log\left[(\hat{Z}_{\mathrm{V@R},2})^{Z_{\mathrm{V@R},1}}(1-\hat{Z}_{\mathrm{V@R},2})^{T-Z_{\mathrm{V@R},1}}\right] - 2\log\left[(\alpha)^{T-Z_{\mathrm{V@R},1}}(1-\alpha)^{Z_{\mathrm{V@R},1}}\right]$; follows asymptotically a chi-squared distribution with 1 degree of freedom $\chi^2(1)$.

Remark: The choice of statistic might change the power of the test. Note, also that although convenient, some care must be taken when using asymptotic results ⚠.

Backtesting ES (unconditional)

Example (Costanzino and Curran 2015) for continuous r.v.:

- Under

$$H_0 : \text{The sequence } \{I_{-t}\}_{t=1,\dots,T} \text{ is i.i.d.,}$$
$$\mathbb{P}[-\Delta V > \mathrm{V@R}^p(\Delta V)] = (1-p), \text{ for all } p \geqslant \alpha,$$

(the whole tail is well-estimated) the estimator

$$\tilde{Z}_{\mathrm{ES},1} = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{1-\alpha}\int_{\alpha}^{1}\mathbb{1}_{\{-\Delta V_{-t}\geqslant \mathrm{V@R}^{u}_{-t}\}}\mathrm{du};$$

is asymptotically normal with mean and variance

$$\mathcal{N}\left(\frac{1-\alpha}{2}, \frac{(1-\alpha)(4+3\alpha)}{12T}\right).$$

Backtesting ES (cont.)

Assume that L follows a continuous distribution. We can use the estimator (Acerbi and Szekely 2014) for continuous r.v.:

- Under

$$H_0 : \mathbb{P}[-\Delta V > \mathrm{V@R}^p(\Delta V)] = (1-p), \text{ for all } p \geqslant \alpha,$$

The estimator

$$Z_{\mathrm{ES},2} = 1 - \sum_{t=1}^{T} \frac{-\Delta V_{-t} I_{-t}}{T(1-\alpha)\mathrm{ES}^{\alpha}_{-t}}.$$

has zero expectation. In this case we do not have an asymptotic convergence result.

To perform the statistical test, it is necessary to run a Monte Carlo simulation based on the full tail distribution available form the null assumption.

Backtesting ES (cont.)

Pseudo-code:

function EMPIRICALSAMPLE_Z(M, T, α, ES^{α})
    $z = \mathrm{ones}(M)$
    for $m = 1$ to M **do**
        for $t = 1$ to T **do**
            $p \sim U[0, 1]$
            if $p \geqslant \alpha$ **then**
                $\Delta v = -\mathrm{V@R}_t^p$
                $z[m] = z[m] + \frac{\Delta v}{T(1-\alpha)ES_t^{\alpha}}$
            end if
        end for
    end for
    return z $\triangleright$ A sample of $Z_{ES,2}$
end function

Testing for clustering

A simple approach to test for clustering is to test for independence of the variables $X_t = I_t$ and $Y = I_{t+1}$. Let

- $n_{i,j} = \#\{t; X_t = i \wedge Y_t = j\}$ for $i, j = 0, 1$.
- $n_{i,.} = \#\{t; X_t = i\}$; $n_{.,j} = \#\{t; Y_t = j\}$

Independence can be checked with a chi-square test of independence:

- H_0 : the variables are independent
- The statistic

$$Q = \sum_{i=0}^{1} \sum_{j=0}^{1} \frac{(n_{i,j} - e)^2}{e},$$

  where $e = \frac{1}{n} n_{i.} n_{.j}$, follows asymptotically a chi-square with 1 degree of freedom.

Comparative testing

Comparative backtest relies on the ellicitability property.

A (law invariant) risk measure ρ is said to be *ellicitable* relative to a class $\mathcal{P}$ of probability measures if there is a scoring function $s : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that

$$\rho(X) = \arg\min_{x \in \mathbb{R}} \mathbb{E}[s(x, X)] \text{ for all } X \text{ with law in } \mathcal{P}.$$

- Some ellicitable risk measures: $\mathrm{V@R}$ and expectiles.
- Some non-ellicitable risk measures: Expected shortfall and standard deviation.

Ellicitability and comparative backtest

To do a comparative backtest we compare an approximation of the scoring function.

Example:
The scoring function of value at risk is

$$s(x, y) = y + \frac{1}{1-\alpha}\mathbb{E}[(x+y)^{-}]].$$

Given a sequence of stationary historical data two approximations of value at risk at level alpha ρ^1, ρ^2, then we compare the largest among the empirical averages for s.

Backtesting ES is harder...

With respect to $\mathrm{V@R}$ backtesting ES requires:

- Stronger null assumptions
- More data to be saved
- More computationally expensive testing procedure
- But unconditional backtesting covering can be done!
- Note: ES is not ellicitable but ellicitability is required for comparative effectiveness but not for unconditional covering.
- However, $\mathrm{V@R}^{\alpha}$ and ES^{α} are jointly ellicitable!

Some additional comments

- There are tests on conditional covering (not assuming independence of the $\mathrm{V@R}$ infringements). See Christoffersen 1998.
- It is convenient to test independently the adequacy of a model. Example: P&L attribution.
- ⚠Backtesting does not solve intrinsic problems with your assumptions.

Aggregation

In many cases, we need to aggregate the risk measure calculation for different types of factors or business lines.

Recall that if the risk measure is coherent (like ES),

$$\rho(X + Y) \leqslant \rho(X) + \rho(Y)$$

Aggregation

In many cases, we need to aggregate the risk measure calculation for different types of factors or business lines.

Recall that if the risk measure is coherent (like ES),

$$\rho(X+Y) \leqslant \rho(X) + \rho(Y)$$

In the case of Gaussian distributions, the aggregation accounts for finding covariance coefficients, since if (X, Y) are jointly Gaussian

$$\mathrm{V@R}^{\alpha}(X+Y) = \sqrt{(\mathrm{var(X)} + \mathrm{var(Y)} + 2\mathrm{cov(X, Y)})}\,\Phi^{-1}(\alpha) - \mathbb{E}(X+Y)$$

Aggregation

In many cases, we need to aggregate the risk measure calculation for different types of factors or business lines.

Recall that if the risk measure is coherent (like ES),

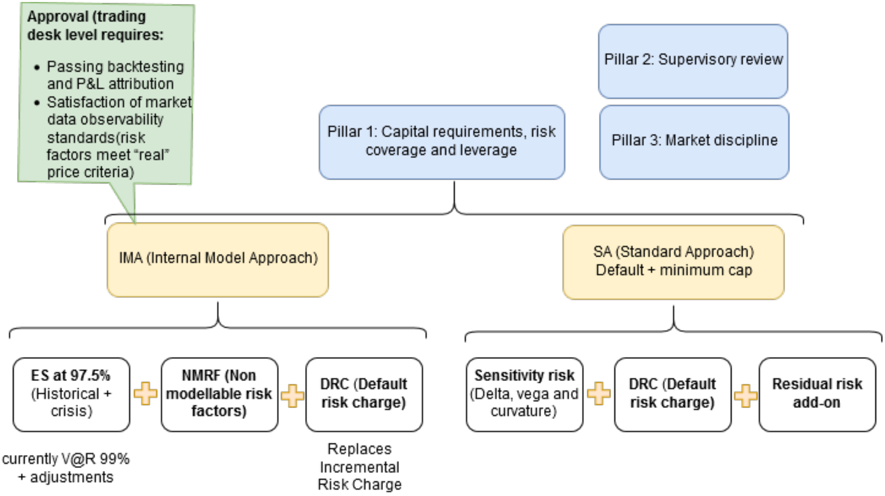
$$\rho(X+Y) \leqslant \rho(X) + \rho(Y)$$

In the case of Gaussian distributions, the aggregation accounts for finding covariance coefficients, since if (X, Y) are jointly Gaussian

$$\mathrm{V@R}^{\alpha}(X+Y) = \sqrt{(\mathrm{var(X) + var(Y) + 2cov(X, Y)})}\,\Phi^{-1}(\alpha) - \mathbb{E}(X+Y)$$

More in general, a *copula* can be used to aggregate different results. See (Nelsen 1999).

Regulation: - Fundamental Review of Trading Book



Standard Approach

- List of risk factors is provided along with associated weights
- Calculate for each position in the trading book and each factor: Delta (sensitivity with respect to price of a factor) and vega (sensitivity with respect to the volatility of the factor); also curvature (an approximation of the gamma).
- Weighted sensitivities with respect to factors in the same 'bucket' are aggregated (by 'Gaussian' rule). Correlations are given
- Totals in between buckets are aggregated (using again 'Gaussian' rule. Correlations are given.

Internal Model Approach

- Risk factors chosen by the bank
- Practical calculation can follow any approach we outlined before and uses Expected shortfall at level 97.5%
- Subject to backtetsing of V@R 99.
- Subject to 'P&L' attribution test: comparing the P&L predicted by the risk model using the factors, with the historical P&L: a good explanation must be found using Spearman test and Kolmogorov-Smirnov.
- Non-modellable factors follow standard approach.

What about designing quantitative investment

The key steps are the same:

i. Choose a factor model
ii. Estimate the criteria for strategy selection and its associated best value (e.g. performance)
iii. (Back) test your results

Additional aspects - investment strategy

- Keep transaction costs low: for example by considering strategies with marginal rebalancing or low rebalancing rate
- Reduce possible market impact: for example by imposing limits on the acceptable positions in given assets
- Avoid survivor bias: by studying performance of assets as chosen when a backtest study starts, and not when it ends
- Robustness is important: include some time regularisation in your criteria.

References I

[AS14] Carlo Acerbi and Balazs Szekely. “Back-Testing Expected Shortfall”. In: *Risk* 27.11 (2014), pp. 76–81.

[CC15] Nick Costanzino and Mike Curran. *Backtesting General Spectral Risk Measures with Application to Expected Shortfall*. SSRN Scholarly Paper ID 2514403. Rochester, NY: Social Science Research Network, Feb. 2015.

[Chr98] Peter F. Christoffersen. “Evaluating Interval Forecasts”. In: *International economic review* (1998), pp. 841–862.

[MFE15] Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton university press, 2015.

[Nel99] Roger B. Nelsen. *An Introduction to Copulas*. Lecture Notes in Statistics 139. New York: Springer, 1999.

[Zie16] Johanna F. Ziegel. “Coherence and Elicitability”. In: *Mathematical Finance* 26.4 (Oct. 2016), pp. 901–918.