Avoiding zero probability events when computing Value at Risk contributions

Rodrigo S. Targino

Escola de Matemática Aplicada (EMAp), FGV

Joint work with Yuri Saporito (EMAp/FGV)

15/Jul/2020
Let us consider $X = (X_1, ..., X_d)$ the losses (negative of the returns) of $d$ different assets in a portfolio.

For a linear portfolio with unitary exposure to each asset, the portfolio-wide loss is defined as $X = \sum_{i=1}^{d} X_i$

Risk measures:

- $\text{VaR}_\alpha(X) = \inf\{x \in \mathbb{R} \mid P(X \leq x) \geq \alpha\}$

- $\text{ES}_\alpha(X) = \frac{1}{1 - \alpha} \int_{\alpha}^{1} \text{VaR}_{1-\beta}(X) d\beta$
Introduction

- After a risk measure of the portfolio is computed, one is usually interested in understanding how much each asset contributes to the overall portfolio risk, in a process known as risk allocation.

- **Euler allocation**: by how much the risk is increased if we increase the exposure to one asset by a small amount.

- Proposed in Tasche (1999) and discussed in several other papers:
  - **Theoretical**: Denault (2001), Kalkbrener (2005), Buch and Dorfleitner (2008)
  - **Empirical**: Tasche (1999), Glasserman (2005), Brownlees and Engle (2012), Mainik and Schaanning (2014), Tasche (2008).


**VaR x ES**

- **Expected Shortfall**
  - Used in the **SST** for capital calculations
  - Fundamental review of the trading book
  - Euler allocations:
    \[ C_i^\alpha = \mathbb{E}[X_i \mid X \geq \text{VaR}_\alpha(X)] \]

- **VaR**
  - Used in **Solvency II** for capital calculations
  - Euler allocations:
    \[ C_i^\alpha = \mathbb{E}[X_i \mid X = \text{VaR}_\alpha(X)] \]

- **VaR x ES**: Embrechts et al. (2014), Emmer et al. (2015)
Objective

- Our aim: To compute Euler allocations for Value at Risk
- Rarely available in closed form
- Given a distribution for $X$, we need to estimate $C_{\alpha}^i = \mathbb{E}[X_i | X = \text{VaR}_\alpha(X)]$
- Exact Monte Carlo: Needs samples from $X | X = \text{VaR}_\alpha(X)$
- Baseline estimator$^1$:
  - Sample from $X | X \in [\text{VaR}_\alpha - \delta(X), \text{VaR}_\alpha + \delta(X)]$
  - $\delta$-allocations: $C_{\alpha,\delta}^i = \mathbb{E}[X_i | X \in [\text{VaR}_\alpha - \delta(X), \text{VaR}_\alpha + \delta(X)]]$

---

$^1$Glasserman (2005)
Literature review

- Exact VaR allocations via Monte Carlo
  - MCMC: Koike and Minami (2019) and Koike and Hofert (2020)
  - Conditional MC: Fu et al. (2009)

- Kernel estimator: Gouriéroux et al. (2000), Tasche (2008), Liu and Hong (2009)

- Infinitesimal perturbation (IPA): Hong (2009)

- Fourier Transform MC: Siller (2013)
▶ Remember ES allocations are easier: \( \mathbb{E}[X_i \mid X \geq \text{VaR}_\alpha(X)] \)

▶ Roudu’s talk\(^2\): Probability Equivalent Level of VaR-ES (PELVE)

▶ The PELVE is the \( c \) such that \( \text{ES}_{1-c\varepsilon}(X) = \text{VaR}_{1-\varepsilon}(X) \)

▶ In principle one could use the allocations for \( \text{ES}_{1-c\varepsilon} \), but it wouldn’t be the same as the allocations for \( \text{VaR}_{1-\varepsilon} \).

---

\(^2\)Li and Wang (2019)
In Asimit et al. (2019) (predecessor to PELVE) the idea is to find $\alpha^*$ such that

$$ES_{\alpha^*}(X) = \text{VaR}_{\alpha}(X)$$

And then compute the $ES_{\alpha^*}(X)$ allocations

For **elliptical distributions**, the allocations based on $ES_{\alpha^*}$ are the same as the allocations based on $\text{VaR}_{\alpha}$

Conditions are provided for when the two allocations are close to each other (for large $\alpha$)
Our proposal

- We also rewrite the VaR allocations as something close to ES allocations

- We identify a model by a function of uniform random variables
  \[ X = g(U) = (g_1(U), \ldots, g_d(U)), \]
  where \( U \sim U[0, 1]^k \) and \( g \in C^1([0, 1]^k; \mathbb{R}^d) \)

- So, we express a \( d \)-dimensional random vector using \( k \)-uniform random variables
Our proposal

**Theorem**

Assume \( X = g(U) \) and that \( \exists f_i \in C^1([0, 1]^k; \mathbb{R}^k) \) s.t., for \( u \in [0, 1]^k \),

\[
\begin{bmatrix}
\nabla g_i(u) \\
\sum_{j \neq i} \nabla g_j(u)
\end{bmatrix} f_i(u) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.
\]

Then, the marginal risk allocation for the VaR is

\[
C_i^\alpha = \frac{\mathbb{E} [X_i \pi_i | X \geq \text{VaR}_\alpha(X)]}{\mathbb{E} [\pi_i | X \geq \text{VaR}_\alpha(X)]},
\]

where \( \pi_i = \text{Tr}(\nabla f_i(U)) \) is the weight, \( \text{Tr}(A) \) is the trace operator of a matrix \( A \) and \( \nabla f_i \) is the Jacobian matrix of \( f_i \).
Our proposal

- The **model** definition is encompassed into the function $g$ (we’ll see examples soon)

- Given a model $g$, one only needs to compute the **weights** $\pi_i$

- Compare this new representation with the ES allocations:

\[
\frac{\mathbb{E}[X_i \mid X \geq \text{VaR}_\alpha(X)]}{\mathbb{E}[\pi_i \mid X \geq \text{VaR}_\alpha(X)]} \quad \text{vs} \quad \mathbb{E}[X_i \mid X \geq \text{VaR}_\alpha(X)]
\]

- Same conditioning events ⇒ **variance reduction** techniques for ES allocations work here as well
  - Targino et al. (2015), Peters et al. (2017), Koike and Minami (2019), Koike and Hofert (2020)
Numerical examples

- We now present the new representation of the VaR allocations for several models.

- We also compare the precision of two Monte Carlo estimators for the allocations.
  1. The new representation $C_i^\alpha = \frac{\mathbb{E}[X_i \pi_i | X \geq \text{VaR}_\alpha(X)\big]}{\mathbb{E}[\pi_i | X \geq \text{VaR}_\alpha(X)\big]}$

  2. The $\delta$-allocation $C_i^{\alpha,\delta} = \mathbb{E}[X_i | X \in [\text{VaR}_{\alpha-\delta}(X), \text{VaR}_{\alpha+\delta}(X)]]$

- We use the same MC sample $X^{(1)}, \ldots, X^{(N)}$ for both methods and a pre-computed VaR.

- We want to empirically assess the impact of $N$ and $\delta$ for $\alpha = 0.5, 0.9$ and 0.99.
Independent marginals

- $X_j = \varphi_j(U_j)$, with $U_1, \ldots, U_d \overset{iid}{\sim} U[0, 1]$ and $\varphi_j$ may be an inverse cdf with differentiable density $p_j$

- Thus, $g_i(u) = \varphi_i(u_i)$ and

$$\nabla g_i(u) = (0, \ldots, 0, \varphi'_i(u_i), 0, \ldots, 0),$$

where the non-zero entry is in the $i$-th position.

- Hence, the following $f$ satisfies the condition in the Theorem

$$f_i(u) = \frac{1}{d-1} \left( \frac{1}{\varphi'_1(u_1)}, \ldots, \frac{1}{\varphi'_{i-1}(u_{i-1})}, 0, \frac{1}{\varphi'_{i+1}(u_{i+1})}, \ldots, \frac{1}{\varphi'_d(u_d)} \right)$$

- Therefore,

$$\text{Tr}(\nabla f_i(u)) = - \frac{1}{d-1} \sum_{j \neq i} \frac{\varphi''_j(u_j)}{(\varphi'_j(u_j))^2} \implies \pi_i = \sum_{j \neq i} \frac{\varphi''_j(U_j)}{(\varphi'_j(U_j))^2} = \sum_{j \neq i} \frac{p'_j(X_j)}{p_j(X_j)}.$$
### Independent marginals

| Name           | Marginal         | $p_j'(x)/p_j(x)$                  |
|----------------|------------------|----------------------------------|
| Log-Normal     | $LN(0, \sigma_j^2)$ | $\frac{Z_j + \sigma_j}{\sigma_j X_j}$, where $Z_j = \frac{1}{\sigma_j} \log X_j$ |
| Exponential    | $Exp(\lambda_j)$  | $\lambda_j$                      |
| Gamma          | $Gamma(\alpha_j, \beta_j)$ | $\left(\frac{\alpha_j - 1}{X_j} - \beta_j \right)$ |
| Gaussian       | $N(0, \sigma_j^2)$ | $-\frac{x}{\sigma_j^2}$          |
| Generalized Pareto | $GPD(\xi_j, \beta_j)$ | $-\frac{1 + \xi_j}{\beta_j} \left(1 + \xi_j \frac{x}{\beta_j}\right)^{-1}$, for $x \geq 0$ |
Independent Log-Normals

- $X_i \overset{ind}{\sim} LN(0, \sigma_i)$
- $\sigma_1 = 0.5$, $\sigma_2 = 1$ and $\sigma_3 = 2$
Independent Log-Normals

**Figure:** Mean (top) and variance (bottom) of the $\delta$-estimator (black) and the new (red) for $\mathcal{C}_1^\alpha$. Line types (solid, dashed and dotted): different values of $\alpha$. Columns: different values for $\delta$. 
Elliptical Distributions

- Elliptical distributions:

\[ X = \mu + RLS \]

- **S** is uniformly distributed in the sphere in \( \mathbb{R}^d \)
- **L** is a \( d \times d \) full-rank, lower triangular matrix
- **R** is an one-dim. radial random variable independent of **S**.

- Gaussian (special case)

\[ X = \mu + LZ. \]
Elliptical Distributions

- The weights for a general elliptical distribution are computed in the preprint.

- **Multivariate Gaussian**: \( \mathbf{X} = \mu + L \mathbf{Z} \)
  - \( \mathbf{Z} \sim N(\mathbf{0}_k, \mathbf{I}_k) \), with \( k \geq d \)
  - \( L \) is a \( d \times k \) full-rank, lower triangular matrix.
  - \( \ell_i \) the \( i \)-th row of the matrix \( L \)

- The weights are \( \pi_i = f_i \cdot \mathbf{Z} \) where \( f_i \) is a solution for
  \[
  \left[ \begin{array}{c} \ell_i \\ \sum_{j \neq i} \ell_j \end{array} \right] \begin{bmatrix} f_i \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}
  \]

- This linear system has infinitely many solutions
Multivariate Gaussian

- \( \mathbf{X} = (X_1, \ldots, X_d) \sim \mathcal{N}(0, \Sigma) \)

- VaR contributions can be computed in closed form
  \[
  \mathcal{C}_{i}^\alpha = \Phi^{-1}(\alpha) \frac{(\Sigma \lambda)_i^T}{\sqrt{\lambda^T \Sigma \lambda}}
  \]

- We also have that \( \text{VaR}_\alpha(\mathbf{X}) = \Phi^{-1}(\alpha)\sqrt{\lambda^T \Sigma \lambda} \).

- For the example:
  - \( \mu = 0 \)
  - \( \Sigma = LL^T \)
  - \( d = 3 \)
  - \( L = \begin{bmatrix} 1 & 0 & 0 \\ 0.5 & 0.7 & 0 \\ 1 & 0.8 & 1.1 \end{bmatrix} \)
  - Variances: 1.0, 0.74 and 2.85
  - Correlations ranging from 0.58 to 0.72.
Figure: Mean (top) and variance (bottom) of the $\delta$-estimator (black) and the new (red) for $\mathcal{C}^\alpha_1$. Line types (solid, dashed and dotted): different values of $\alpha$. Columns: different values for $\delta$. 
Archimedean copulas

- $C_\psi$ is an Archimedean copula with generator $\psi$ and $X = (X_1, \ldots, X_d)$ has joint cdf

$$F(x_1, \ldots, x_d) = C_\psi(F_1(x_1), \ldots, F_d(x_d))$$

- To generate one sample from $X$ we
  1. Sample $V \sim F = LS^{-1}(\psi)$
  2. Sample $U_i \overset{iid}{\sim} U[0, 1], i = 1, \ldots, d$
  3. Define $U_i = \psi(-\log(U_i)/V), i = 1, \ldots, d$
  4. Define $X_i = F_i^{-1}(U_i)$

- Notation/hypothesis:
  - $C_\psi$ is an Archimedean copula with generator $\psi$
  - $F = LS^{-1}(\psi)$ the inverse Laplace-Stieltjes transform of $\psi$
  - Both $F$ and the marginals $F_i$ are absolutely continuous
  - $p_i$ (density of $F_i$) is differentiable
Archimedean copulas

Notation:

- $\mathcal{H} = \mathcal{F}^{-1}$
- $\phi_j(u) = -\log(u_j)/\mathcal{H}(u_k)$
- $\gamma_j(u) = \frac{\mathcal{H}(u_k)}{\psi' (\phi_j(u))} + \frac{\psi'' (\phi_j(u))}{\psi' (\phi_j(u))^2}$

For Archimedean copulas the weights are given by

$$\pi_i = \sum_{j \neq i, k} p_j(X_j) \gamma_j(U) - \frac{p'_j(X_j)}{p_j(X_j)}$$

For survival Archimedean copulas,

$$\pi_i = \sum_{j \neq i, k} p_j(X_j) \gamma_j(U) + \frac{p'_j(X_j)}{p_j(X_j)},$$
### Archimedean copulas

| Copula | $\gamma_j(u)$ |
|--------|---------------|
| **Clayton** | $\psi(t) = (1 + t)^{-1/\vartheta}$, $\mathcal{V} \sim \Gamma(1/\vartheta, 1)$ \[ \frac{1}{\psi(\phi_j(U))} (-\vartheta(\mathcal{V} - \log U_j) + \vartheta + 1) \] |
| **Gumbel** | $\psi(t) = e^{-t^{1/\vartheta}}$, $\mathcal{V} \sim S\left(\frac{1}{\vartheta}, 1, c, 0; 1\right)$ \[ \frac{1}{\psi(\phi_j(U))} \left(-\vartheta \mathcal{V}\phi_j(U)^{1-1/\vartheta} + (\vartheta - 1)\phi_j(U)^{-1/\vartheta} + 1 \right) \] |
Survival Clayton with GPD marginals

- Survival Clayton copula with parameter $\theta = 2$
- Kendall’s tau $\tau = 0.5$
- $d = 3$
- $X_i \sim GPD(\xi_i, \beta_j)$
- $\xi_i = 0.3$ (moments up to order 3 are finite)
- $\beta_i = 1$

<Model M1 from Koike and Hofert (2020)>
Survival Clayton with GPD marginals

**Figure:** Mean (top) and variance (bottom) of the $\delta$-estimator (black) and the new (red) for $C_1^\alpha$. Line types (solid, dashed and dotted): different values of $\alpha$. Columns: different values for $\delta$. 
Conclusions

- We are able to derive a novel expression for the Value-at-Risk contributions.

- We go from an expectation conditional to a zero probability event in the usual representation, to a ratio of expectations conditional to events of positive probability.

- The new formulation is amenable to Monte Carlo simulation with mild hypothesis on the multivariate models and the precise formulas are provided for a wide range of models.

- The new representation shows promising results when compared to a simple estimator.

- As the expectations in the proposed formulation resemble the Expected Shortfall allocations from which algorithms could be adapted for further computational gains.
Do we have time for Math?

YES! :)  no :(  

Appendix

- The main theorem was presented for models of the form $X = g(U)$.

- Without loss of generality, we abuse the notation and discuss the proof when $X = g(Z)$.

- The proof uses Malliavin calculus.

- A less technical proof using only integration by parts may also be possible.

- We explain later why we decided to use Malliavin calculus instead of integration by parts.
Malliavin calculus is a differential calculus for functionals of the **Brownian motion**

**Notation:**

- \((W_t)_{t \in [0, T]} : k\)-dimensional Brownian motion,
- \(W_t = (W^1_t, \ldots, W^k_t)\)
- \((\mathcal{F}_t)_{t \in [0, T]}\) the filtration generated by \((W_t)_t\)
- \(\mathbb{D}^{1,2} : \) space of r.v.’s in \(L^2(\Omega, \mathcal{F}_T, \mathbb{P})\) that are differentiable in the Malliavin sense
Appendix

- A very important subspace of $\mathbb{D}^{1,2}$ is the space of smooth random variables

$$F = g \left( \int_0^T h_1(s) dW_s, \ldots, \int_0^T h_n(s) dW_s \right),$$

with $g \in C_c^\infty(\mathbb{R}^n)$ and $h \in L^2([0, T]; \mathbb{R}^k)$.

- In this case, the Malliavin derivative at time $t \leq T$, which is denoted by $D_t$, is given by

$$D_t F = \sum_{k=1}^n \partial_k g \left( \int_0^T h_1(s) dW_s, \ldots, \int_0^T h_n(s) dW_s \right) h_k(t),$$

where $\partial_k g$ is the derivative of $g$ with respect to the $k$th variable.
An important case for our application is $F = g(W_1^T, \ldots, W_k^T)$, where $g \in C^1(\mathbb{R}^k)$.

In this case,

$$D_tF = \nabla g(W_1^T, \ldots, W_k^T)$$

In the multivariate case where $F = (F^1, \ldots, F^m)$, the Malliavin derivative $D_tF$ is a $m \times k$ matrix where the $j$th row is given by $D_tF^j$. 

The adjoint operator of $D$, denoted by $\delta$ and called **Skorokhod integral**, is defined by the integration-by-parts formula:

$$
\mathbb{E}[F \delta(v)] = \mathbb{E} \left[ \int_0^T D_t F \cdot v_t dt \right], \quad \forall \ F \in \mathbb{D}^{1,2}
$$

The domain of $\delta$ is characterized by the $\mathbb{R}^k$-valued stochastic processes $v = (v_t)_{t \in [0, T]}$ (**not necessarily adapted** to the filtration $(\mathcal{F}_t)_{t \in [0, T]}$) such that

$$
\left| \mathbb{E} \left[ \int_0^T D_t F \cdot v_t dt \right] \right| \leq C \|F\|_2, \quad \forall \ F \in \mathbb{D}^{1,2},
$$

where $C > 0$ might depend on $v$ and $\|F\|_2 = \mathbb{E}[|F|^2]^{1/2}$.
Important: For $F_j$ a smooth random variable and $h_j \in L^2([0, T]; \mathbb{R}^k)$, $j = 1, \ldots, m$,

$$\delta \left( \sum_{j=1}^{m} F_j h_j \right) = \sum_{j=1}^{m} \left( F_j \int_{0}^{T} h_j(t) dW_t - \int_{0}^{T} D_t F_j \cdot h_j(t) dt \right).$$

For smooth r.v.’s the Skorohod integral can be computed in terms of Ito and Riemman integrals.
Appendix

The cornerstone of our result is the following theorems from Ewald (2005) and Fournié et al. (2001)
Appendix

Theorem

Let $F, G \in \mathbb{D}^{1,2}$ such that $F$ is $\mathbb{R}^m$-valued, $G$ is $\mathbb{R}$-valued with $D_t G$ non-degenerate. Assume there exists a process $v$ in the domain of $\delta$ and

$$
\mathbb{E} \left[ \int_0^T D_t G \cdot v_t dt \, | \, F, G \right] = 1.
$$

Assume further that $\phi \in C^1(\mathbb{R})$. Then

$$
\mathbb{E}[\phi(F) \, | \, G = 0] = \frac{\mathbb{E} \left[ \phi(F) \delta(v) H(G) - \phi'(F) H(G) \int_0^T D_t F v_t dt \right]}{\mathbb{E}[\delta(v) H(G)]},
$$

where $H(x) = 1_{x \geq 0}$ is the Heaviside function.
Additionally to the assumptions of the theorem above, assume

\[ \mathbb{E} \left[ \int_0^T D_t F \cdot \nu_t \, dt \ \Big| \ F, G \right] = 0_m, \]

where \( 0_m \) is the \( m \)-dimensional zero vector. Then, for any Borel measurable function \( \phi \) with at most linear growth at infinity,

\[ \mathbb{E}[\phi(F) \mid G = 0] = \frac{\mathbb{E}[\phi(F)\delta(\nu)H(G)]}{\mathbb{E}[\delta(\nu)H(G)]}. \]
Thank you for your attention!
Asimit, V., Peng, L., Wang, R., and Yu, A. (2019). An efficient approach to quantile capital allocation and sensitivity analysis. Mathematical Finance, 29(4):1131–1156.

Brownlees, C. T. and Engle, R. (2012). Volatility, correlation and tails for systemic risk measurement. Available at SSRN, 1611229.

Buch, A. and Dorfleitner, G. (2008). Coherent risk measures, coherent capital allocations and the gradient allocation principle. Insurance: Mathematics and Economics, 42(1):235–242.

Denault, M. (2001). Coherent allocation of risk capital. Journal of Risk, 4(1):1–34.

Embrechts, P., Puccetti, G., Rüschendorf, L., Wang, R., and Beleraj, A. (2014). An academic response to Basel 3.5. Risks, 2(1):25–48.

Emmer, S., Kratz, M., and Tasche, D. (2015). What is the best risk measure in practice? A comparison of standard measures. Journal of Risk, 18(2):31–60.

Ewald, C.-O. (2005). Local volatility in the Heston model: a Malliavin calculus approach. Journal of Applied Mathematics and Stochastic Analysis, 2005(3):307–322.
Fournié, E., Lasry, J.-M., Lebuchoux, J., and Lions, P.-L. (2001). Applications of Malliavin calculus to Monte-Carlo methods in finance. II. *Finance and Stochastics*, 5(2).

Fu, M. C., Hong, L. J., and Hu, J.-Q. (2009). Conditional monte carlo estimation of quantile sensitivities. *Management Science*, 55(12):2019–2027.

Glasserman, P. (2005). Measuring marginal risk contributions in credit portfolios. *Journal of Computational Finance*, 9(2):1.

Gouriéroux, C., Laurent, J. P., and Scaillet, O. (2000). Sensitivity analysis of values at risk. *Journal of Empirical Finance*, 7(3):225–245.

Hong, L. J. (2009). Estimating quantile sensitivities. *Operations research*, 57(1):118–130.

Kalkbrener, M. (2005). An axiomatic approach to capital allocation. *Mathematical Finance*, 15(3):425–437.

Koike, T. and Hofert, M. (2020). Markov chain monte carlo methods for estimating systemic risk allocations. *Risks*, 8(1):6.

Koike, T. and Minami, M. (2019). Estimation of risk contributions with mcmc. *Quantitative Finance*, 19(9):1579–1597.

Li, H. and Wang, R. (2019). Pelve: Probability equivalent level of var and es. Available at SSRN.
Liu, G. and Hong, L. J. (2009). Kernel estimation of quantile sensitivities. Naval Research Logistics (NRL), 56(6):511–525.

Mainik, G. and Schaanning, E. (2014). On dependence consistency of covar and some other systemic risk measures. Statistics & Risk Modeling, 31(1):49–77.

Peters, G. W., Targino, R. S., and Wüthrich, M. V. (2017). Bayesian modelling, monte carlo sampling and capital allocation of insurance risks. Risks, 5(4):53.

Siller, T. (2013). Measuring marginal risk contributions in credit portfolios. Quantitative Finance, 13(12):1915–1923.

Targino, R. S., Peters, G. W., and Shevchenko, P. V. (2015). Sequential Monte Carlo samplers for capital allocation under copula-dependent risk models. Insurance: Mathematics and Economics, 61:206–226.

Tasche, D. (1999). Risk contributions and performance measurement. Report of the Lehrstuhl für mathematische Statistik, TU München.

Tasche, D. (2008). Capital allocation to business units and sub-portfolios: the Euler principle. In Pillar II in the New Basel Accord: The Challenge of Economic Capital, pages 423–453. Risk Books.