Public

Extreme Value Theory and Application

A mission improbable

Alessio Brussino
Statistical Risk Aggregation Methodology

May 20, 2019

SECTION 1

Theory in a nutshell

UBS

Extreme Values

Definition

t-years event: given a random variable *X* with distribution function *F*, for *t* sufficiently large, is the event corresponding to a value

$$x_t = F^{\leftarrow}(1 - t^{-1})$$

where $F^{\leftarrow}$ is the generalized inverse of *F*.

Goal

Calculation of the *upper-quantile q* = 1 - t^{-1} of the df *F*, where *t* is larger than the lenght of the time-series of the available empirical observations (eg. *t* = 100-10'000 years), i.e. the level that with probability *q* will not be superceeded.

Applications

- Engineering, hydrology (1953 North Sea flood, 1953-98 Delta Works, 2004-2011 tsunami)
- Geology, meteorology, environmental analysis
- Insurance (fire & storm claims, reinsurance, …)
- **Finance (market & operational risk, stress testing)**

Historical Background

*"… In the case of the dyke problem, what were the available historical data? The basic data consisted of observations on maximum annual heights. Moreover, the previous record surge occurred in 1570 with a recorded value of (NAP+4)m, where NAP stands for normal Amsterdam level. The 1953 flooding corresponded to a (NAP+3.85)m surge. The statistical analysis in the Delta report led to an **estimated (1 - t^{-1})-quantile of (NAP+5.14)m** for the yearly maximum, based on a total of 166 historical observations.(…) Clearly, the task of estimating such a 10000year event leads to estimating well beyond the range of the data."*

Out-of-sample Estimation

Estimation well beyond the range of the data: the solution requires making extra assumptions on the underlying model.

UBS

Useful Tools

Empirical distribution function $$F_n(x) = \frac{\#\{i : 1 \le i \le n \quad and \quad X_i \le x\}}{n}, \quad x \in \Re$$

Order Statistics $$X_{1,n} = \max(X_1, \ldots, X_n) \ge X_{2,n} \ge \ldots \ge X_{n,n} = \min(X_1, \ldots, X_n)$$

p-quantile

$$x_p = F^{\leftarrow}(p) = \inf\{x \in \Re : F(x) \ge p\}$$

$$x_{p,n} = F_n^{\leftarrow}(p) = X_{k,n} \qquad 1 - \frac{k}{n} < p \le 1 - \frac{k-1}{n}$$

Shortfall distribution $$P(X - u_L \le -x \mid X < u_L) = \vartheta(x \mid u_L)$$

Excess distribution *F_u* $$F_u(x) = P(X - u \le x \mid X > u), \quad x \ge 0$$

Graphical Exploratory Analysis

Mean Excess Function (MEF): the mean value of the excesses above a certain threshold *u*. It is useful to classify the tail of the distribution (short vs long-tailed).

Theoretical MEF: $e(u) = E(X - u \mid X > u), \quad u \geq 0$

Empirical MEF:
$$e(u) = \frac{\sum_{i=1}^{n} (X_i - u)^+}{\sum_{i=1}^{n} I_{\{X_i > u\}}}$$

QQ-Plot: used to test the goodness-of-fit of a parametric model $F(., \theta)$ compared to the empirical distribution of the values X_i. For appropriate plotting sequence $p_{k,n}$ (in particular $X_{k,n}$):

$$\left\{\left(X_{k,n}, \hat{F}^{\leftarrow}(p_{k,n})\right) : k = 1, \ldots, n\right\}$$

Mean Excess Function: Graphical Example

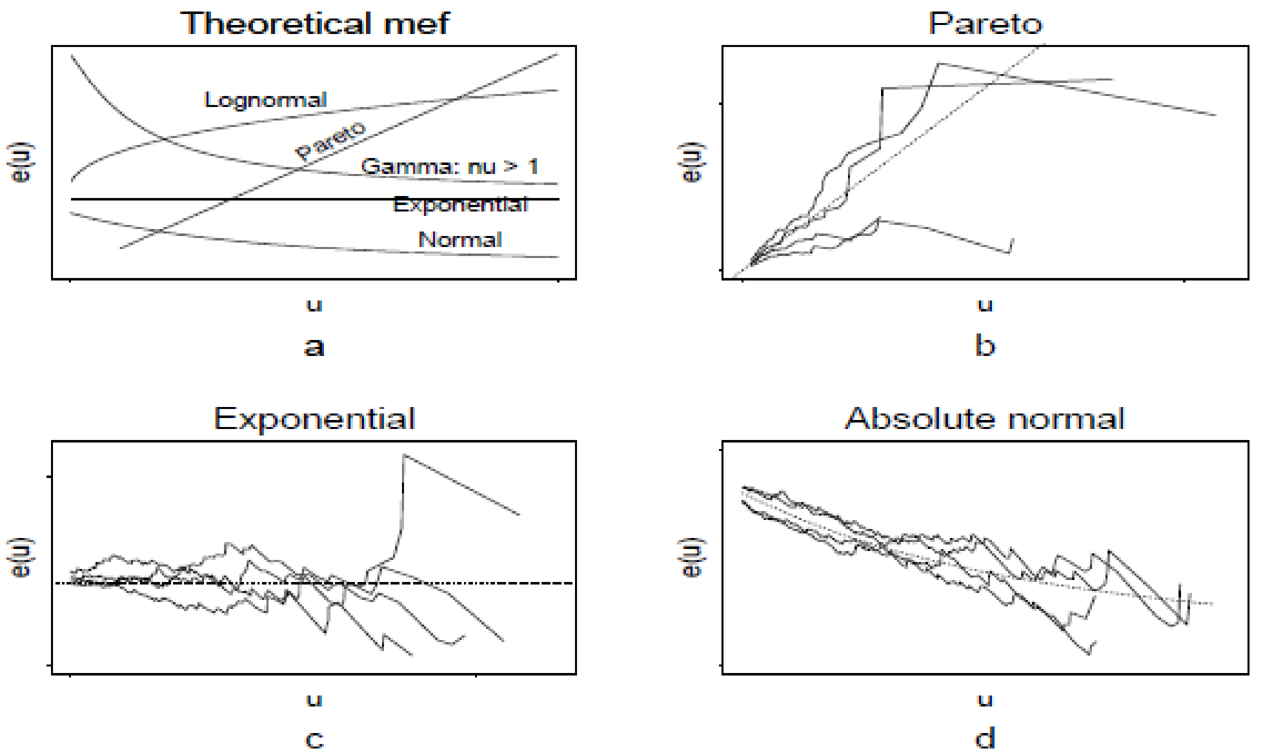


Figure 5: *Theoretical mefs (Figure a) and some examples of empirical mefs for simulated data (n=200). Normal in Figure a stands for the absolute value of a normally distributed rv. In Figure b the empirical mef of 4 simulated Pareto*(1.5) *distributed data-sets are plotted, Figure c corresponds to Exp*(1)*, and d to the absolute value of a normally distributed rv. The dotted line shows the theoretical mef of the underlying distribution.*

Sum Statistics vs Maximum Statistics

Suppose $X_1,\ldots,X_n$ are iid with df F.

Define the **partial sum** $\boldsymbol{S_n = X_1 + \ldots + X_n}$.

If there exist constants $a_n>0$ and b_n real so that:

$$\frac{S_n - b_n}{a_n} \xrightarrow{d} Y \quad , \quad n \to \infty$$

where Y is non-degenerate with df G.

<u>Special Case:</u> distributions with finite variance

then $a_n = \sqrt{n}\sigma$, $b_n = n\mu$ and $G = N(0,1)$

All distributions with finite second moment are "attracted" to the standard normal.

This is the **<u>Central Limit Theorem (CLT)</u>**

Suppose $X_1,\ldots,X_n$ are iid with df F.

Define the **maximum** $\boldsymbol{M_n = X_{1,n} = \max(X_1,\ldots,X_n)}$.

If there exist constants $a_n>0$ and b_n real so that:

$$\frac{M_n - b_n}{a_n} \xrightarrow{d} Y \quad , \quad n \to \infty$$

where Y is non-degenerate with df G, then G is one of the following types:

1. Gumbel

$$\Lambda(x) = \exp\left\{-e^{-x}\right\}$$

2. Fréchet

$$\Phi_\alpha(x) = \begin{cases} 0 & if \quad x \le 0 \\ \exp\left\{-x^{-\alpha}\right\} & if \quad x > 0 \end{cases} \quad for \quad \alpha > 0$$

3. Weibull

$$\Psi_\alpha(x) = \begin{cases} \exp\left\{-(-x)^{\alpha}\right\} & if \quad x < 0 \\ 1 & if \quad x \ge 0 \end{cases} \quad for \quad \alpha > 0$$

This is the **<u>Fisher-Tippett Theorem (FT)</u>**

Generalized Extreme Value Distributions (GEV)

These three classes of dfs can be reparametrised using a GEV df characterized by the parameters: $\xi \in \Re$ *(shape),* $\beta > 0$ *(scale),* $\mu \in \Re$ *(location):*

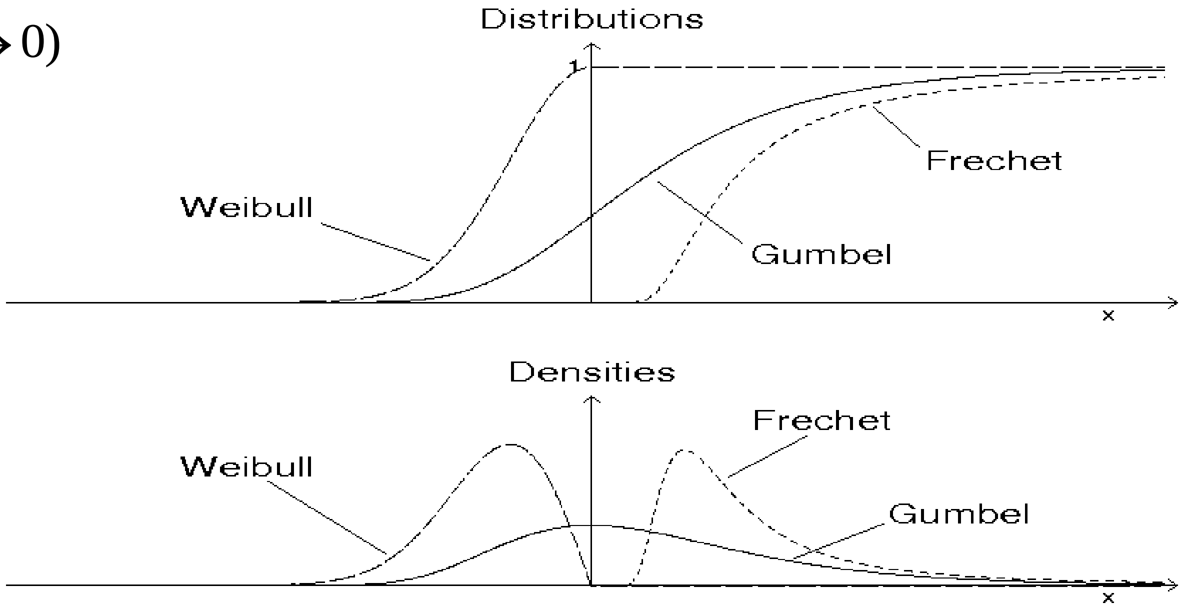
$$H_{\xi,\beta,\mu}(x) = \exp\left\{ -\left(1+\xi \frac{x-\mu}{\beta}\right)^{-\frac{1}{\xi}} \right\}, \quad 1+\xi \frac{x-\mu}{\beta} \geq 0$$

Particular values of ξ correspond to the case:

- Gumbel: $"\xi = 0" \quad (\xi \to 0)$

- Fréchet: $\xi > 0$

- Weibull: $\xi < 0$



UBS

Block Maxima Method

In these terms the problem of the ***maximum domain of attraction* (MDA)** is formalized:

the solution is given by dfs F such that, given H_ξ, FT theorem holds for appropriate sequences a_n and b_n. In this case we can say $\boldsymbol{F \in MDA(H_\xi)}$: F belongs to the MDA of H_ξ with norming constants a_n, b_n.

Suppose $X_1,\ldots,X_n$ are iid with df $F \in MDA(H_\xi)$, then:

$$P(a_n^{-1}(M_n - b_n) \le x) = P(M_n \le a_n x + b_n) = F^n(a_n x + b_n) \;\rightarrow\; H_\xi(x), \qquad n \rightarrow \infty$$

where $u = a_n x + b_n$. Using Taylor expansion one can obtain an *upper-tail* estimator:

$$\hat{\bar{F}}(u) = \frac{1}{n}\left(1 + \hat{\xi}\,\frac{u - \hat{b}_n}{\hat{a}_n}\right)^{-\frac{1}{\xi}}$$

And then an estimator for the quantile x_p is available:

$$\boxed{\hat{x}_p = \hat{b}_n + \frac{\hat{a}_n}{\hat{\xi}}\left(\left(n(1-p)\right)^{-\hat{\xi}} - 1\right)}$$

UBS

Tail Index Estimators

Hill estimator:

$$\hat{\xi}_{n,k}^{H} = \frac{1}{k}\sum_{j=1}^{k} \ln X_{j,n} - \ln X_{k,n}$$

Pickands estimator:

$$\hat{\xi}_{n,k}^{P} = \frac{1}{\ln 2} \ln \frac{X_{k,n} - X_{2k,n}}{X_{2k,n} - X_{4k,n}}$$

Dekkers – Einmahl - de Haan estimator:

$$\hat{\xi}_{n}^{D} = 1 + H_n^{(1)} + \frac{1}{2}\left(\frac{(H_n^{(1)})^2}{H_n^{(2)}} - 1 \right)^{-1}$$

$$H_n^{(1)} = \frac{1}{k}\sum_{j=1}^{k} (\ln X_{j,n} - \ln X_{k+1,n}) \qquad H_n^{(2)} = \frac{1}{k}\sum_{j=1}^{k} (\ln X_{j,n} - \ln X_{k+1,n})^2$$

Gnedenko – Pickands – Balkema - de Haan Theorem

Suppose F is a df with excess distribution F_u, $u \geq 0$. Then, for $\xi \in \Re$, $F \in MDA(H_\xi)$ if and only if there exists a positive measurable function $\beta(u)$ so that

$$\lim_{u \to x_F} \sup_{0 \leq x \leq x_F - u} \left| \overline{F}_u(x) - \overline{G}_{\xi,\beta(u)}(x) \right| = 0$$

Where G is a **Generalized Pareto Distribution** with parameters $\beta > 0$, $\xi \in \Re$ so defined:

$$G_{\xi,\beta}(x) = \begin{cases} 1 - \left(1 + \xi \dfrac{x}{\beta}\right)^{-\frac{1}{\xi}}, & \xi \neq 0 \\ 1 - \exp\left\{-\dfrac{x}{\beta}\right\}, & \xi = 0 \end{cases}$$

Shape parameter

Scale parameter

GPD Properties:

- ξ **> 0** : Pareto distribution (heavy-tailed)
- ξ **= 0** : Exponential distribution
- ξ **< 0** : Pareto type II (short-tailed)

$$\boxed{G_{\xi,\beta}(x) \in MDA(H_\xi) \qquad \forall \xi \in \Re}$$

$$\boxed{\lim_{\xi \to 0} G_{\xi,\beta}(x) = G_{0,\beta}(x)}$$

$$\boxed{if \quad \xi > 0, \quad E[X^k] = \infty \quad \Leftrightarrow \quad k \geq \frac{1}{\xi} \qquad \Rightarrow \qquad \exists \quad \mu = E[X] = \frac{\beta}{1-\xi} \quad \Leftrightarrow \xi < 1}$$

UBS

Peaks-Over-Threshold Method

As a consequence, it is possible to use an *upper-tail estimator* for $x,u \geq 0$ (in particular for $u \to \infty$):

$$\overline{F}(u+x) = \overline{F}_u(x)\overline{F}(u) \approx \overline{G}_{\hat{\xi},\hat{\beta}}(x)\frac{N_u}{n}$$

Where the GPD parameters can be obtained through estimation (eg. MLE), while N_u:

$$N_u = \#\{i : 1 \leq i \leq n, X_i > u\}$$

Finally it is possible to evaluate the quantile x_p:

$$\hat{x}_p = u + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{n}{N_u}(1-p)\right)^{-\hat{\xi}} - 1\right) \qquad ES_p = \frac{VaR_p}{1-\xi} + \frac{\beta - \xi u}{1-\xi}$$

UBS

SECTION 2

ETV Applications: Operational Risk Modeling

UBS

Operational risk: regulatory definition

From Solvency II

Operational risk *is the risk of a change in value caused by the fact that actual losses, incurred for inadequate or failed internal processes, people and systems, or from external events (including legal risk), differ from the expected losses.*

From the technical point of view:

- (1) The operational-risk measure is a VaR at confidence level 99.9 % with a holding period of one year.
- (2) The measurement approach must capture potentially severe tail loss events.

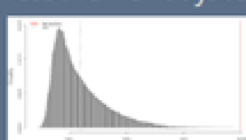
Operational risk: examples

Basel categories and taxonomies

#	Event Type	Examples
1	Internal Fraud	misappropriation of assets, tax evasion, intentional mismarking of positions, bribery
2	External Fraud	theft of information, hacking damage, third-party theft and forgery
3	Employment Practices and Workplace Safety	discrimination, workers compensation, employee health and safety
4	Clients, Products, and Business Practice	market manipulation, antitrust, improper trade, product defects, fiduciary breaches, account churning
5	Damage to Physical Assets	natural disasters, terrorism, vandalism
6	Business Disruption and Systems Failures	utility disruptions, software failures, hardware failures
7	Execution, Delivery, and Process Management	data entry errors, accounting errors, failed mandatory reporting, negligent loss of client assets

Operational Risk Capital calculation

Methods for determining the minimum capital requirement

Approach	Basis	Factor	Result
Basic Indicator Approach (BIA)	Average gross Income of the last 3 years	x 15%	Minimum Capital Requirement
Standardised Approach (SA)	Retail Banking Retail Brokerage Asset Management (Gross Income)	x 12%	Minimum Capital Requirement
	Commercial Banking Agency Services (Gross Income)	x 15%	
	Corporate Finance Trading and Sales Payment & Settlement (Gross Income)	x 18%	
Advanced Measurement Approach (AMA)	Operational Risk Measurement System 	Internal Model*	Minimum Capital Requirement

Less

Qualitative Qualifying Standards

More

UBS

AMA Capital calculation

Building Blocks of an Operational Risk Quantification Approach

Internal Losses	External Losses	Scenario Analysis	Business Environment and Internal Control Factors (BEICF)

According to the BCBS Supervisory Guidelines, an AMA framework must include the use of four data elements: (i) Internal loss data (ILD), (ii) External data (ED), (iii) Scenario analysis (SBA), and (iv) Business environment and internal control factors (BEICFs).

The banking industry has developed different types of models to estimate operational risk that include the four key elements in different ways.

We will focus on the category of models defined by the regulator as **LDA**:

> *Under the Loss Distribution Approach, the bank estimates, for each business line/risk type cell, the probability distribution functions of the single event impact and the event frequency for the next (one) year using its internal data, and computes the probability distribution function of the cumulative operational loss.*

UBS

LDA: Loss Distribution Approach

Model component	Description	Modelling choice
Units of Measure (UoM) = Risk Category	External Fraud, Market Conduct, etc.	• UoM taxonomy to satisfy homogeneity, not necessarily corresponding to Basel categories
Event Frequency (per UoM)	N_j: Random number of losses per year for UoM j	• Usually Poisson or Negative Binomial distribution
Event Severity (per UoM)	X_i: Random size of a loss event i	• Body: Empirical distribution or Continuous version of the empirical distribution • Tail: parametric "heavy tailed" distribution (Lognormal, Truncated lognormal, Pareto, GPD, Weibull, Burr, etc.)

Cumulative loss per year (per UoM)	$Y = X_1 + X_2 + \cdots + X_{N\_j}$
	
	Convolution of frequency and severity distributions
		• Calculation of cumulative loss distribution (CLD) per UoM via:
		    • Deterministic algorithms: FFT, Panjer, SLA
		    • Monte Carlo algorithms

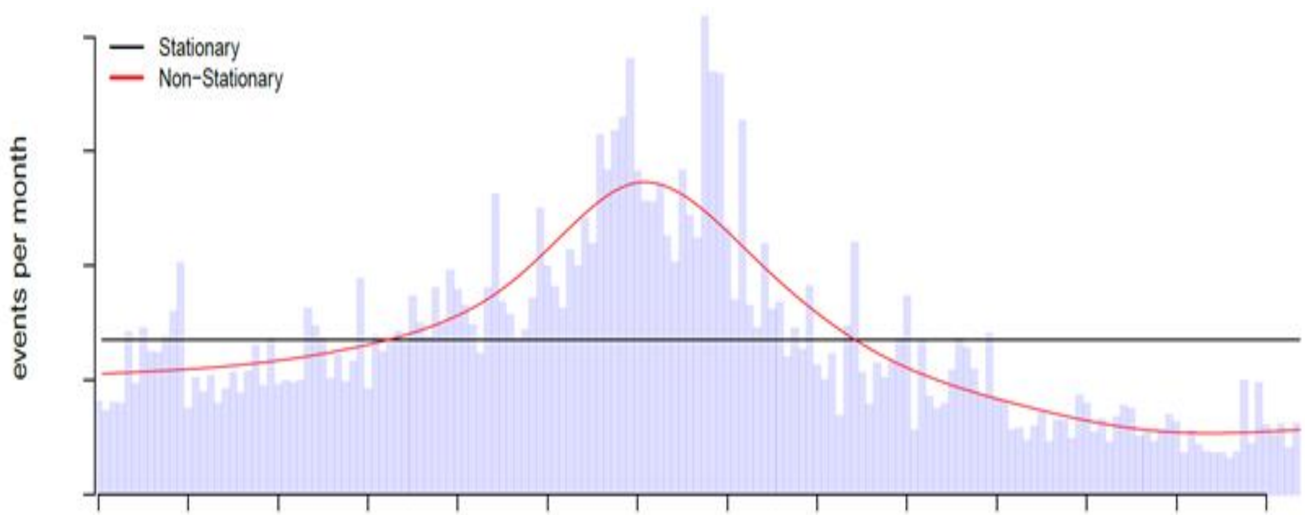
Diversification/Dependence between UoM cumulative loss distributions	To aggregate the cumulative losses per UoM to the total loss, a model for the dependence structure between cumulative losses per UoM is needed	• Independence is usually not conservative • Copula (Gaussian, t, asymmetrical) • "perfect dependence": comonotonicity
Capital calculation	Calculation of the normative capital requirement as the quantile (99.9%) of the total loss distribution	• Aggregation of UoM CLD's via copula using Monte Carlo simulation. To control Monte Carlo simulation error, capital calculation can be repeated. • Calculation of 99.9% quantile of the total yearly losses

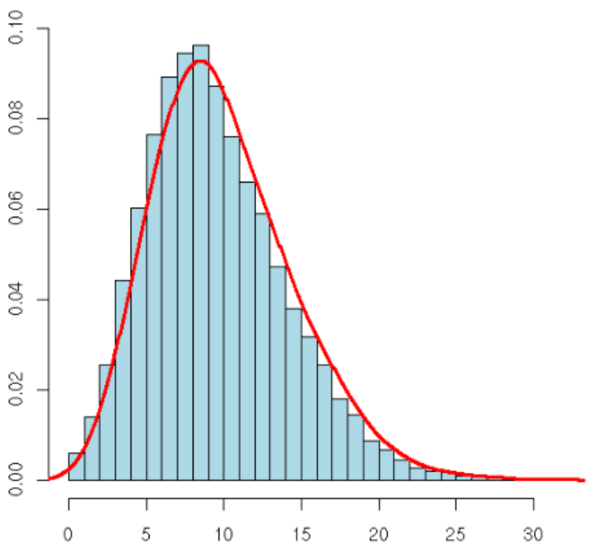
Frequency Estimation

Distribution of the number of events per UoM per year

Key elements to consider:

- Independence between frequency and severity distribution is an assumption
- Presence of temporary evolution (stationary vs non-stationary models)
- Presence of under/overdispersion (Poisson, negBin, etc.)



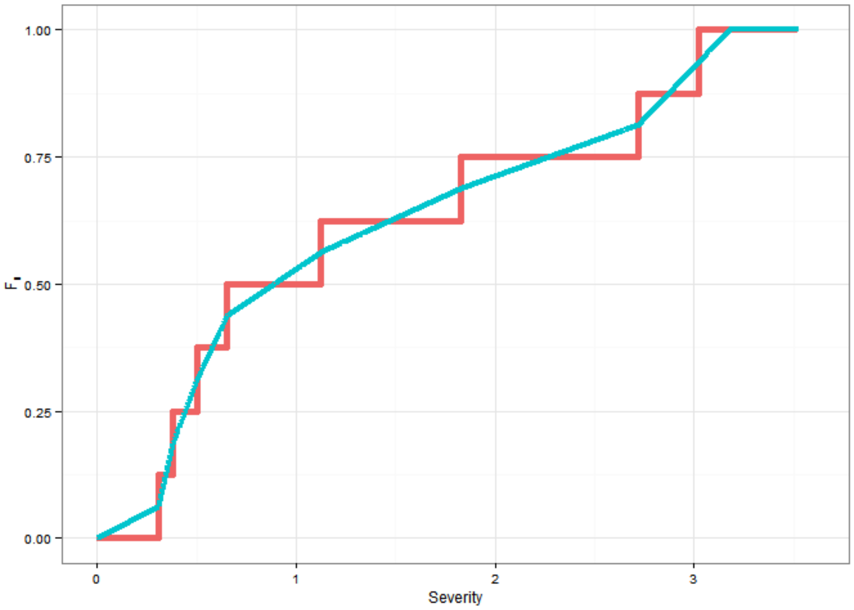


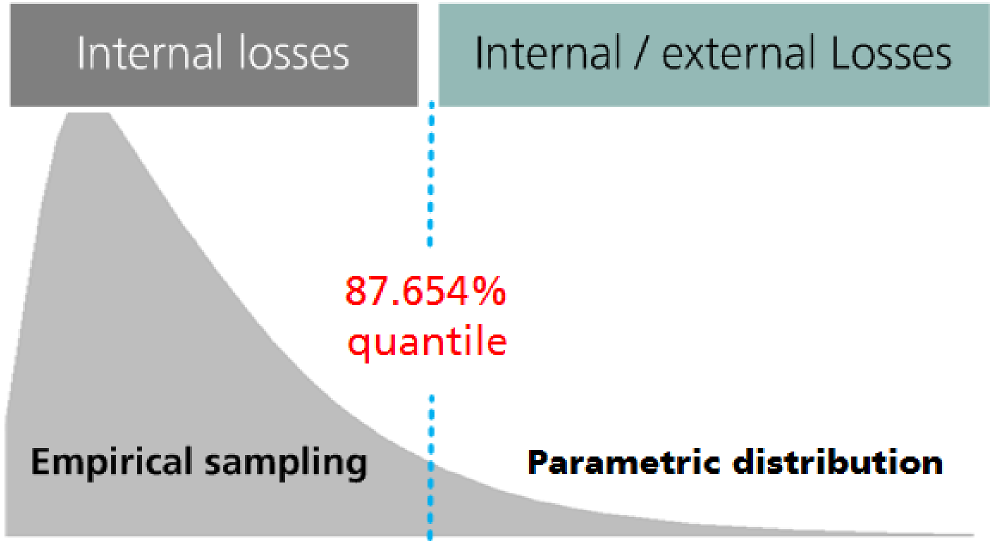
Severity Estimation

Distribution of the size of the single events per UoM

Key elements to consider:

- Independence between the size of different events in the severity distribution is an assumption.
- Use of a mixed distribution (body + tail) to better represent the empirical properties.
- The parametric piece with heavy tailed distr. to account for (potentially Inf.) events not historically experienced is a regulatory requirement.
- EVT distributions are a common choice to fit the parametric tail.





UBS

SLA: Single Loss Approximation

Operational VaR: a Closed-Form Approximation

Klaus Böcker Claudia Klüppelberg

Abstract

Klaus Böcker and Claudia Klüppelberg investigate a simple loss distribution model for operational risk. They show that, when loss data are heavy-tailed (which in practice they are), a simple closed-form approximation for the OpVaR can be obtained. They apply this approximation in particular to the Pareto severity model, for which they obtain also a simple time scaling rule for the operational VaR.

For rv X_i with subexponential distribution F:

$$\lim_{x\to\infty} \frac{P(X_1+\cdots+X_n>x)}{P(\max(X_1,\ldots,X_n)>x)} = 1 \quad \text{for some (all) } n \geq 2.$$

Under mild conditions for the frequency distribution – satisfied by Poisson, negBin – it can be proved that for the tail of the aggregate loss distribution G_t:

$$\overline{G}_t(x) \sim EN(t)\overline{F}(x)\,, \quad x \to \infty\,, \tag{2}$$

where EN is the expected frequency.

It follows an analytical solution of the convolution problem for asymptotic quantiles:

Theorem 2.4 (Analytical OpVaR). *Consider the Standard LDA model for fixed $t > 0$ and a subexponential severity with distribution function F. Assume, moreover, that the tail estimate (2) holds. Then, the $VaR_t(\kappa)$ satisfies the approximation*

$$VaR_t(\kappa) = F^{\leftarrow}\left(1 - \frac{1-\kappa}{EN(t)}(1+o(1))\right), \quad \kappa \to 1\,. \tag{3}$$

UBS

References

L. de Haan, Fighting the arch-enemy with mathematics. *Statistica Neerlandica* **44** 1990.

F.Bassi, P. Embrechts, M. Kafetzaki, A survival kit on quantile estimation. *ETH, Zurich* (1993).

P. Embrechts, C. Kluppelberg, T. Mikosch, Modelling Extremal Events for Insurance and Finance. *Springer, Berlin* (2003).

A.J. McNeil, R. Frey, P. Embrechts, Quantitative Risk Management. *Princeton University Press, Princeton-NJ* (2005)

Thanks for your attention!

Contacts:

alessio.brussino@ubs.com

Public

Risk Aggregation

Tecniques & Tools

Alessio Brussino
Statistical Risk Aggregation Methodology

May 27, 2019

Section 1

A portfolio perspective of risk

UBS

What is Risk Management

- Process of identifying sources of risk, quantifying risk and developing strategies to manage risk.
- The focus is on risk that result in negative consequences for economic activity.
- Conventional risk management deal with physical or legal risks like natural disasters, accidents, lawsuits.
- Our focus in this class in on financial risk management.

Why is Risk Management necessary?

- To protect against extreme losses.
- To be used as a management tool.
- To strengthen the stability of the financial system (regulation).
- To protect investors and shareholders values.

Why Risk Measurement?

- Determination of risk capital: determine the amount of capital a financial institution needs to cover unexpected losses.
- Management tool: risk measures are used by management to limit the amount of risk a unit within the firm may take.
- Insurance premiums: compensation to insurance companies for bearing risks are based on measures of risk of the insured claims.
- Elementary measures of risk: factor sensitivity measures, scenario based risk measures (stress testing), risk measures based on the loss distribution.

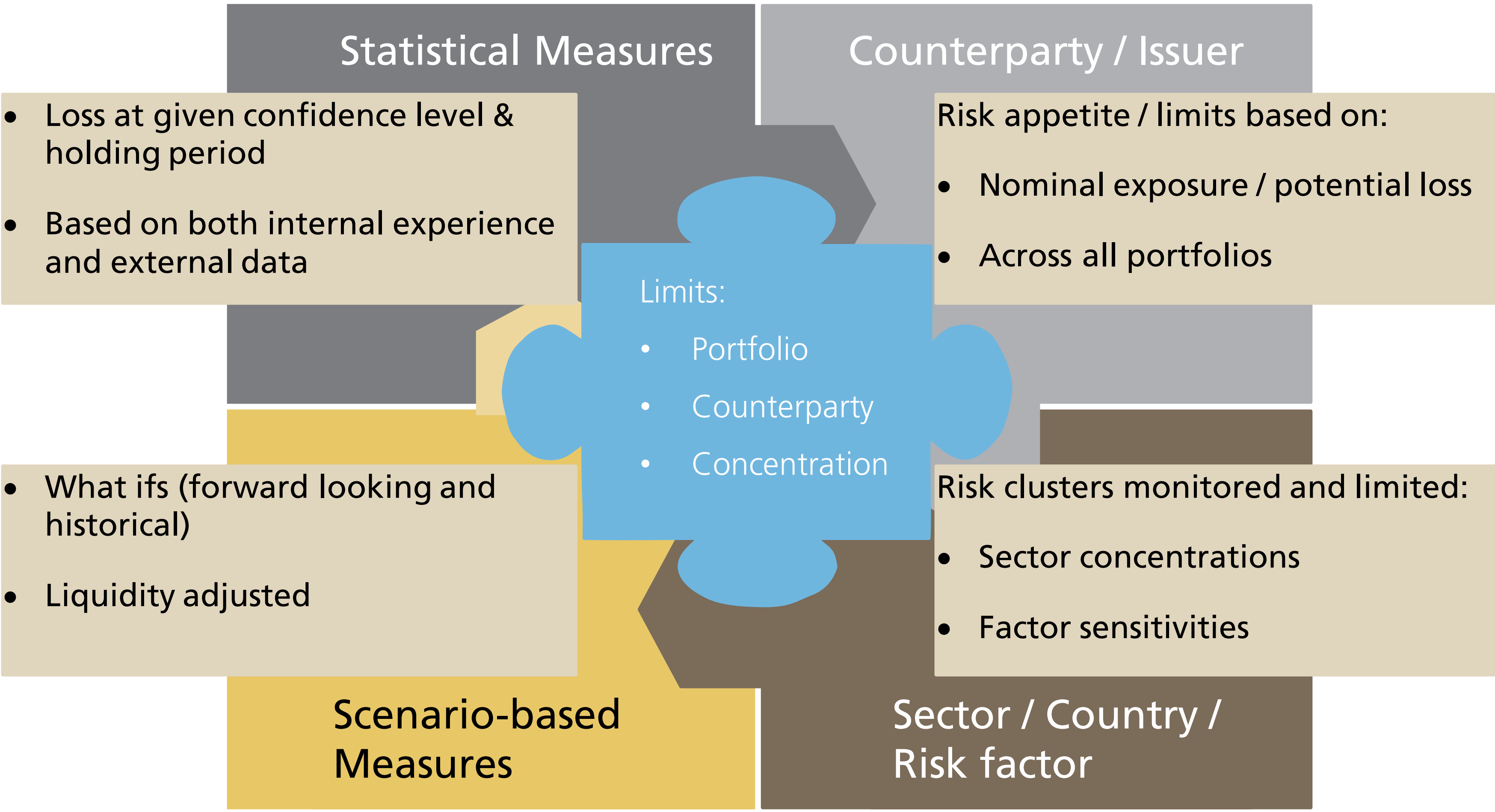
Risk categories

Primary Risks	Consequential Risks	Business Risk
• Market Risk – Investment Risk – Issuer Risk • Credit Risk – Settlement Risk – Loan Underwriting Risk • Country Risk	• Operational Risk • Liquidity and Funding Risk • Pension Risk • Structural FX Risk • Environmental and Social Risk	• Business environment • Economic cycle • Industry cycle • Strategy
Reputational Risk		

Source: based on UBS Group AG annual report 2015, section 3: Risk, treasury and capital management

Risk measurement and control

Application of robust models and rigorous identification of risks



UBS

Portfolio risk measures

The two measures are largely complementary.

Statistical loss

- Derive the distribution of potential earning based on historical observations;

Pros:

- Can be linked to probability/confidence level;
- Can aggregated based on statistical technique;

Cons:

- Results are not intuitive;

Stress testing

- What if:
  - Forward looking: US recession, China hard landing, Grexit, etc
  - Historical severe events: Eurozone crisis, subprime crisis, etc

Pros:

- Built-in links from macro variables to risk factors;
- Results are more intuitive;

Cons:

- Difficult to assign probability to the events;

Statistical risk measure

Mathematical formulation

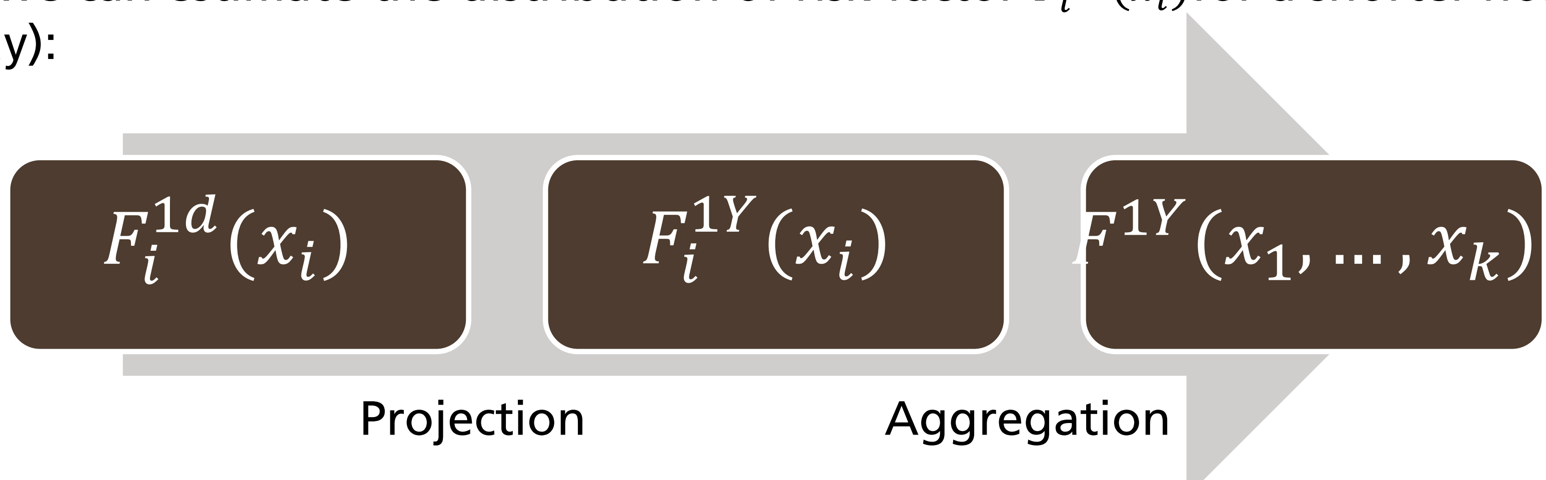
- PnL is known given the realization of risk factors: $L = H(x_1, \dots, x_k)$
- The joint distribution of risk factors, uniquely specified by its cumulative distribution function:

$$F(x_1, \dots, x_k) = P(X_1 \leq x_i, \dots, X_k \leq x_k)$$

The marginal distribution of risk factor X_i is

$$F_i(x_i) = P(X_i \leq x_i) = F(x_i, \infty, \dots, \infty)$$

- Statistical risk measure (VaR, ES, etc) can be evaluated when we know $F(x_1, \dots, x_k)$.
- The question we are facing is forecasting the distribution in given investment horizon (1 year, 1 month, etc) F^{1Y}.
- Suppose we can estimate the distribution of risk factor $F_i^{1d}(x_i)$for a shorter horizon (e.g. 1 day):



Section 2

Risk measures

P&L distribution

- Let V_t be the value of the portfolio now (time t).
- Δ is the fixed time horizon. Typically: one day (Δ = 1/250), ten days, one month or one year.
- V_{t+1} - V_t = $V_{t+\Delta}$ - V_t is the profit over the time interval of length delta and is a random variable.
- The distribution of V_{t+1} - V_t is called the profit-and-loss distribution (P&L).

Loss distribution

- Risk management concentrate on losses.
- It is standard convention to work with the right tail of the distribution.
- L_{t+1} = -(V_{t+1} - V_t) = -($V_{t+\Delta}$ - V_t) is the loss over one time interval of length Δ.
- L_{t+1} will be a random variable whose distribution is called *loss distribution*.

Risk measures based on loss distribution

The risk is measured by summary statistics of the loss distribution

- Standard deviation
- Value-at-Risk (VaR)
- Expected shortfall (ES)
- Advantage: loss distribution can be compared across potfolios with the same time horizon delta an it reflects diversification.
- Disadvantage: loss distribution is estimated by past data

Risk measures based on loss distribution

Standard deviation

- Profit and losses have equal impact on the standard deviation.
- No information on how large potential losses may be.
- Does not discriminate between distribution with different probability of potential large losses.

Risk measures based on loss distribution

Value-at-Risk

Maximum loss which is not exceeded with a given high probability.

$$
\begin{aligned}
VaR_\alpha &= \inf\{l \in \mathbb{R}: P(L > l) \le 1 - \alpha\} \\
&= \inf\{l \in \mathbb{R}: F_L(l) \le 1 - \alpha\} \\
&= q_\alpha
\end{aligned}
$$

- $F_L(l) = P(L \le l)$ distribution function of the loss L.
- Time horizon Δ (one day, ten days or one year).
- Confidence level $\alpha \in (0,1)$ (95%, 99%, 99.9%).
- Negative diversification effects can arise (non-subadditivity).
- VaR not provide information by how much any actual will exceed VaR.

Risk measures based on loss distribution

Value-at-Risk for normal distribution

- Assume $L \sim N(\mu, \sigma^2)$

$$P\{L \leq l\} = P\left\{\frac{L - \mu}{\sigma} \leq \frac{l - \mu}{\sigma}\right\} = \Phi(\frac{l - \mu}{\sigma})$$

- $VaR_{\alpha}(L) = q_{\alpha}(F_L)$

$$P\{L \leq VaR_{\alpha}(L)\} = \Phi\left(\frac{VaR_{\alpha}(L) - \mu}{\sigma}\right) = \alpha$$

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

Risk measures based on loss distribution

Expected shortfall

$$ES_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 VaR_u(L)\,du$$

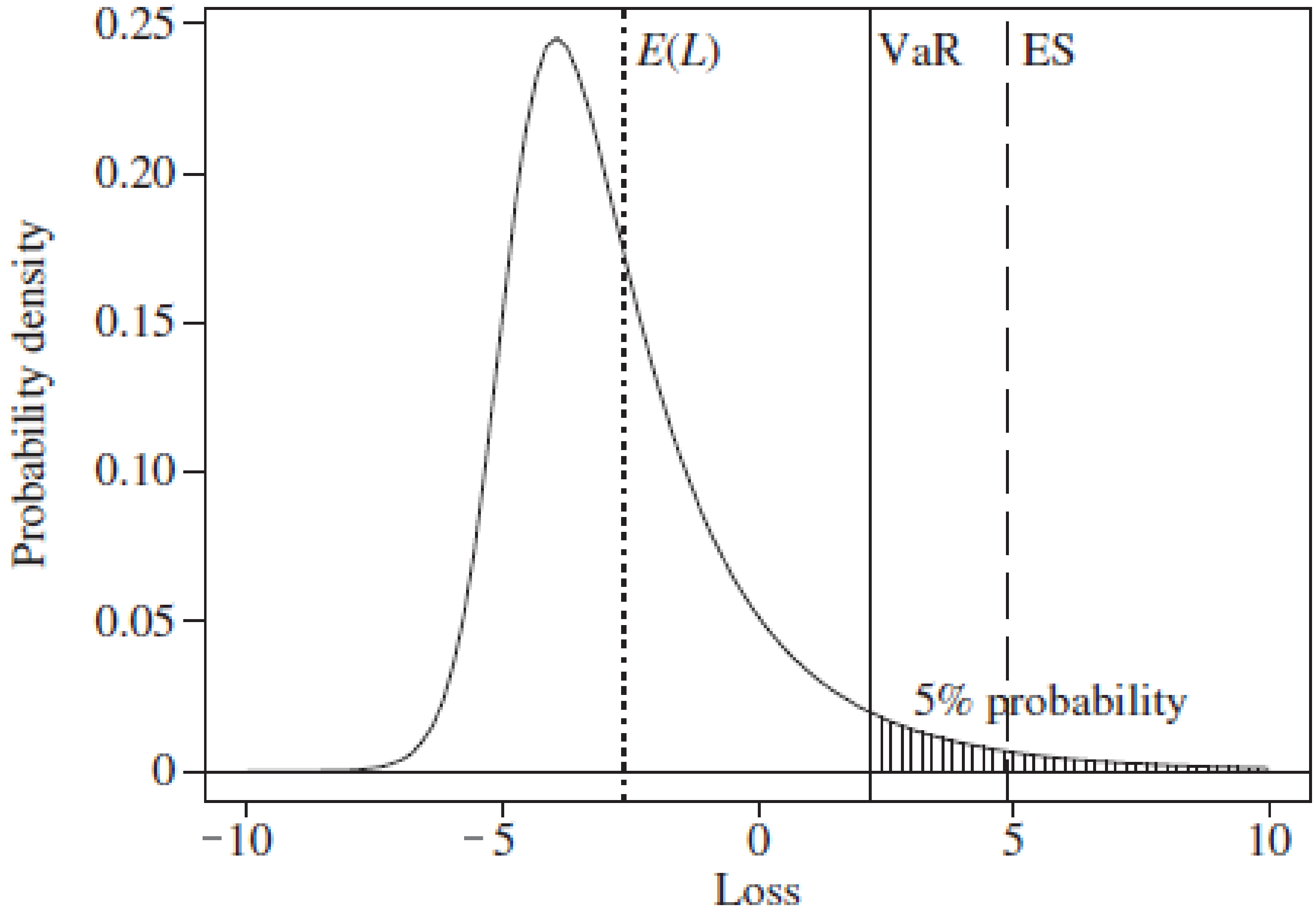
- Confidence level $\alpha \in (0,1)$

If distribution $F_L(l)$ is continuous then expected shortfall is the expected loss that is incurred in the event that VaR is exceeded.

$$ES_\alpha = \mathrm{E}(\mathrm{L} \mid \mathrm{L} \geq VaR_\alpha(L))$$

Risk measures based on loss distribution

Example of loss distribution with the 95% VaR marked as a vertical line, and dotted line an alternative risk measure know as the 95% expected shortfall.



Risk measures based on loss distribution

Example: VaR and Es for stock returns

- We consider daily losses on a position in a particular stock; the current value of the position equals to V_t = 10’000.
- The loss of this portfolio is given by $L_{t+1} = -V_t X_{t+1}$, where X_{t+1} represents daily log-returns of the stock.
- We assume that X_{t+1} has zero mean and standard deviation $\sigma = \frac{0.20}{\sqrt{250}}$, i.e. we assume that the stock has an annualized volatility of 20%.

- TASK: Compare two different models for the distribution:

1. Normal distribution
2. t distribution with $\nu = 4$ degrees of freedom scaled to have standard deviation σ

Risk measures based on loss distribution

VaR_α and ES_α for normal and t model for different values of α

α	0.90	0.95	0.975	0.99	0.995
VaR_α (normal model)	162.1	208.1	247.9	294.3	325.8
VaR_α (t model)	137.1	190.7	248.3	335.1	411.8
ES_α (normal model)	222.0	260.9	295.7	337.2	365.8
ES_α (t model)	223.4	286.3	356.7	465.8	563.5

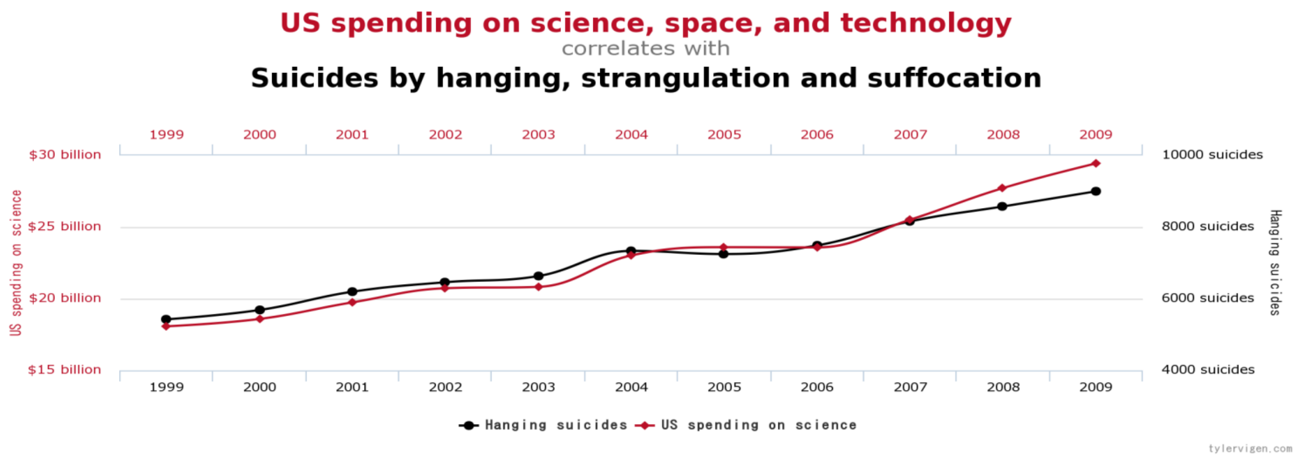
Section 3

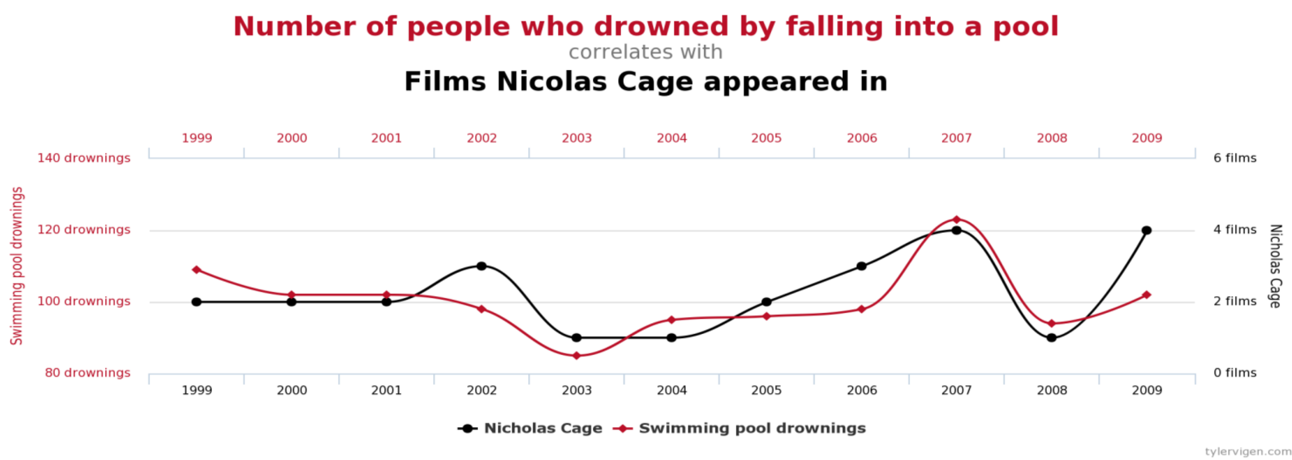
Correlation and dependence

UBS

Some examples of funny "correlation"

Before we head into another technical section





Source: tylervigen.com

Example

The role of dependence in a portfolio

PnL: Scenario I

Year	Portfolio A	Portfolio B	Total
1	-5	12	7
2	10	15	25
3	7	20	27
4	-1	4	3
5	14	2	16
6	27	33	60
7	-10	-19	-29
8	34	-7	27
9	6	11	17

Example

PnL: Scenario II

Year	Portfolio A	Portfolio B	Total
1	-5	12	7
2	10	33	43
3	7	4	11
4	-1	-19	-20
5	14	11	25
6	27	20	47
7	-10	15	5
8	34	-7	27
9	6	2	8

Measure of correlation

- Ideally a function with value from -1 to 1;
  - 0: uncorrelated;
  - 1: perfect positive correlation;
  - -1: perfect negative correlation;
- Pearson's linear correlation coefficient;
- Spearman's Rho
- Kendall's Tau

Example

Inter-correlation between risk types

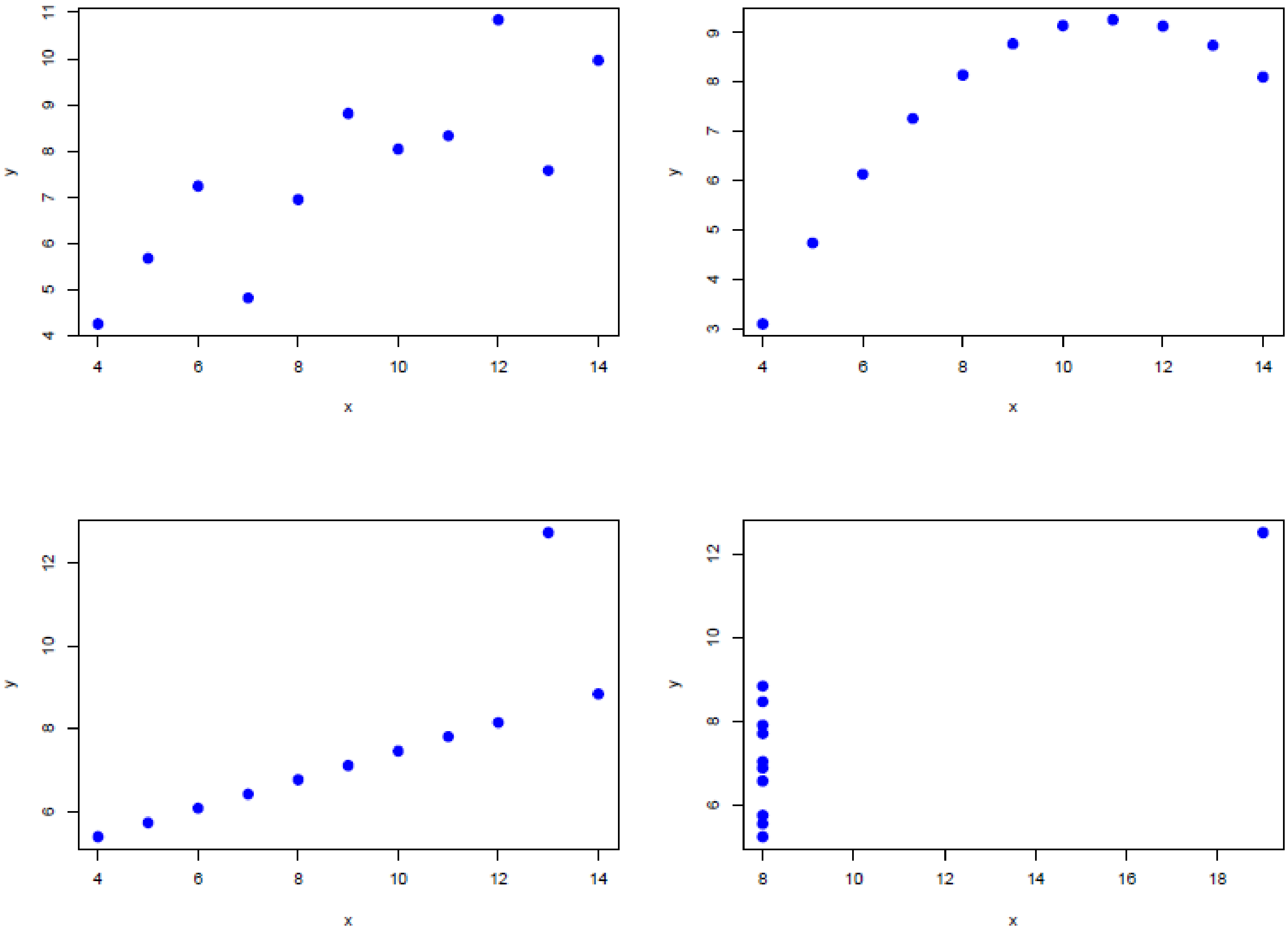
- Market and credit risk (MR/CR) with operational risk (OR):
  - OR leads to increase in MR/CR: Knight Capital Group, 911
  - MR/CR drives OR events: Enron
- Choice of measure?

  - focus on the assumption you make.

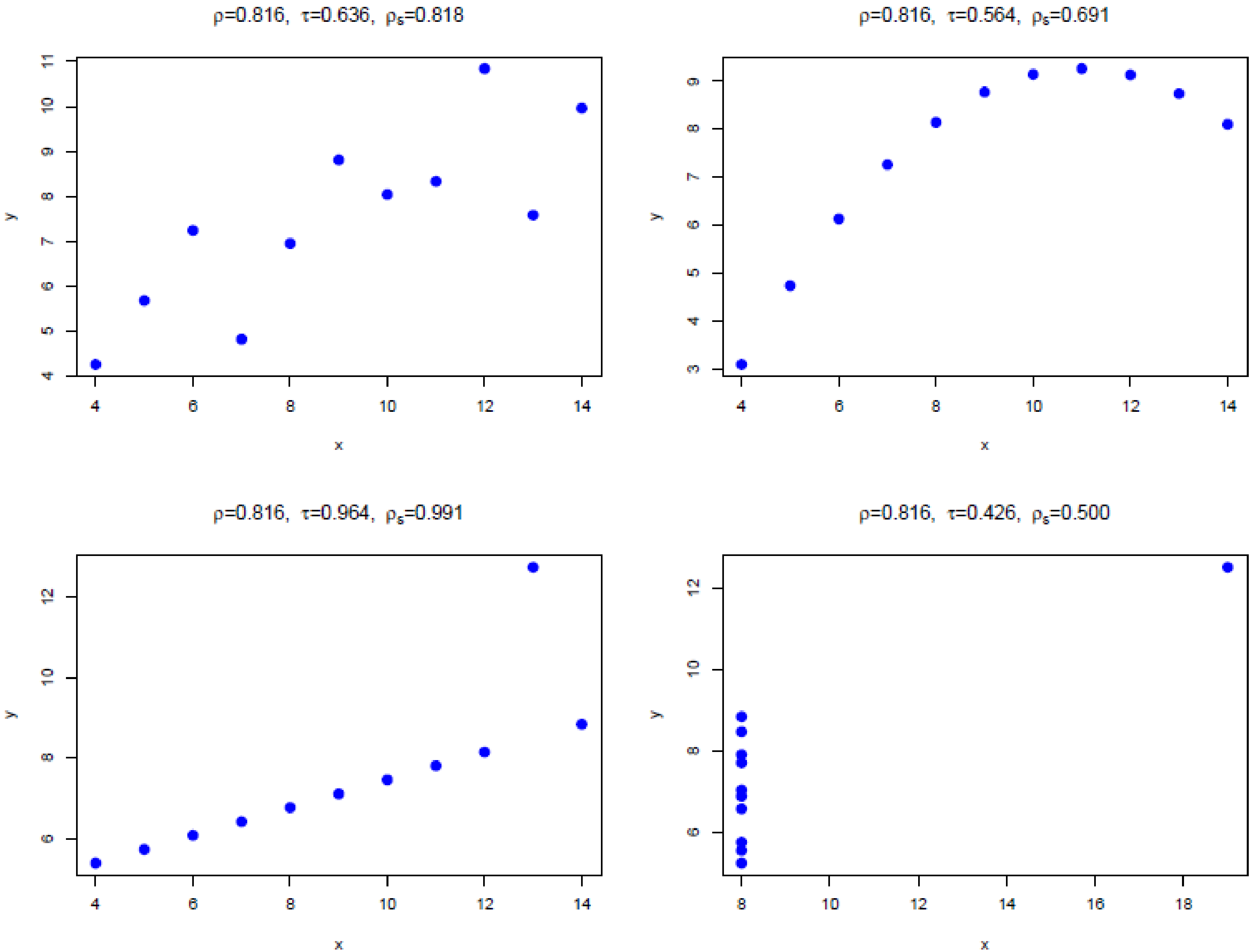
  - Non parametric V.S. parametric

  - Does your data have outliers?

Example



Example: Anscombe's quartet data



Section 4

Copula approach to portfolio risk

UBS

Aggregation - industrial practice

Typically used aggregation methodologies

	Non linearity	Parametrization
Simple summation	✗	Simple
Fixed diversification percentage	✗	Simple
Variance-covariance matrix	✗	Moderate
Copula	✓	Difficult
Full simulation of common risk drivers	✓	Difficult

Base on: BCBS March 2009: Range of practices and issues in economic capital frameworks

UBS

What is copula?

- Sklar's theorem: *There exists a unique cumulative distribution function* $C(u_1, \dots, u_k)$ *on the unit hypercube* $[0,1]^k$ *with standard uniform marginals such that the following representation holds true:*

$$F(x_1, \dots, x_k) = C(F_1(x_1), \dots., F_k(x_k))$$

- For absolute continuous F and strictly increasing, continuous marginal $F_1, \dots, F_k$, we have:

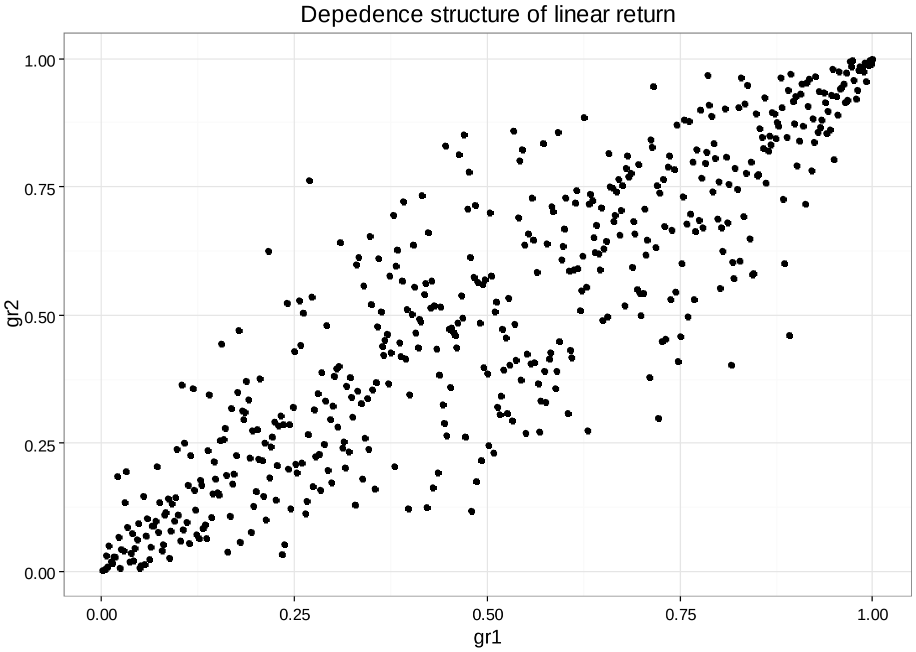
$$f(x_1, \dots, x_k) = c(F_1(x_1), \dots, F_k(x_k)) f_1(x_1) \dots f_k(x_k) \qquad (1)$$

- Copula measures the joint feature of multivariate distribution;
  - One can separately model the marginal distributions and the dependence structure.
  - The copula is invariant under monotonic transformations.
- Example: stock 1 and 2, the log return follow bivariate normal distribution with correlation 0.9

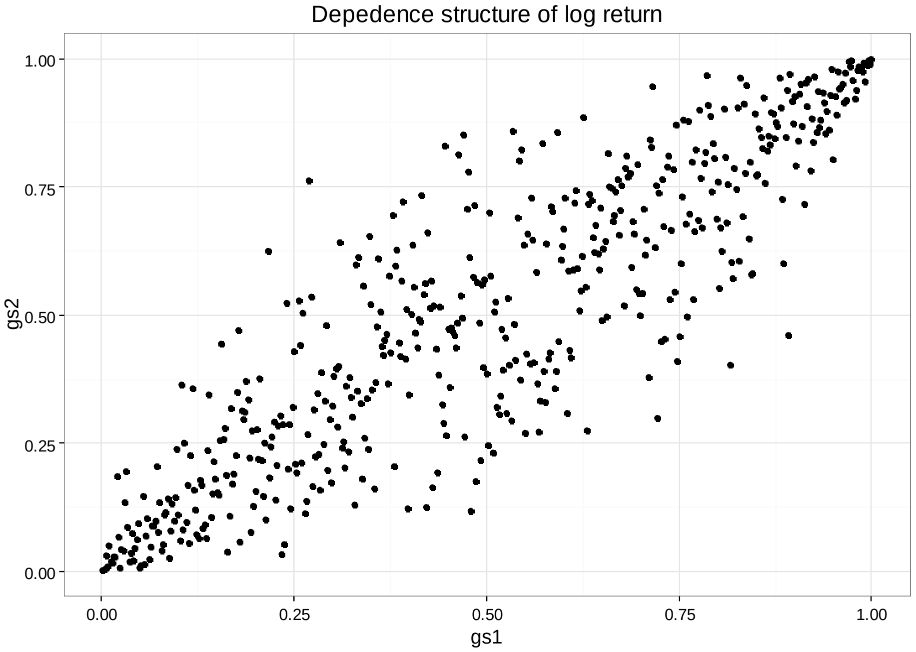
What is copula?

Invariant under monotonic transformation

Copula: linear return

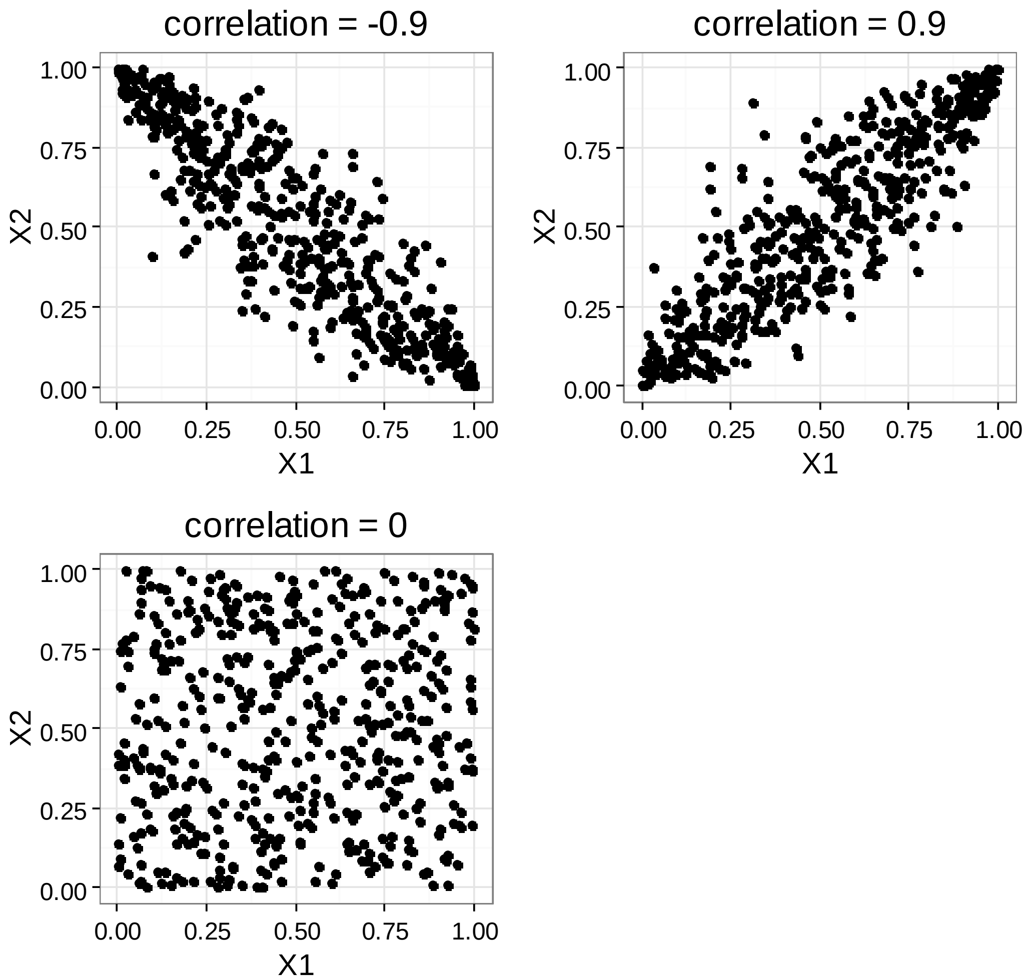


Copula: log return



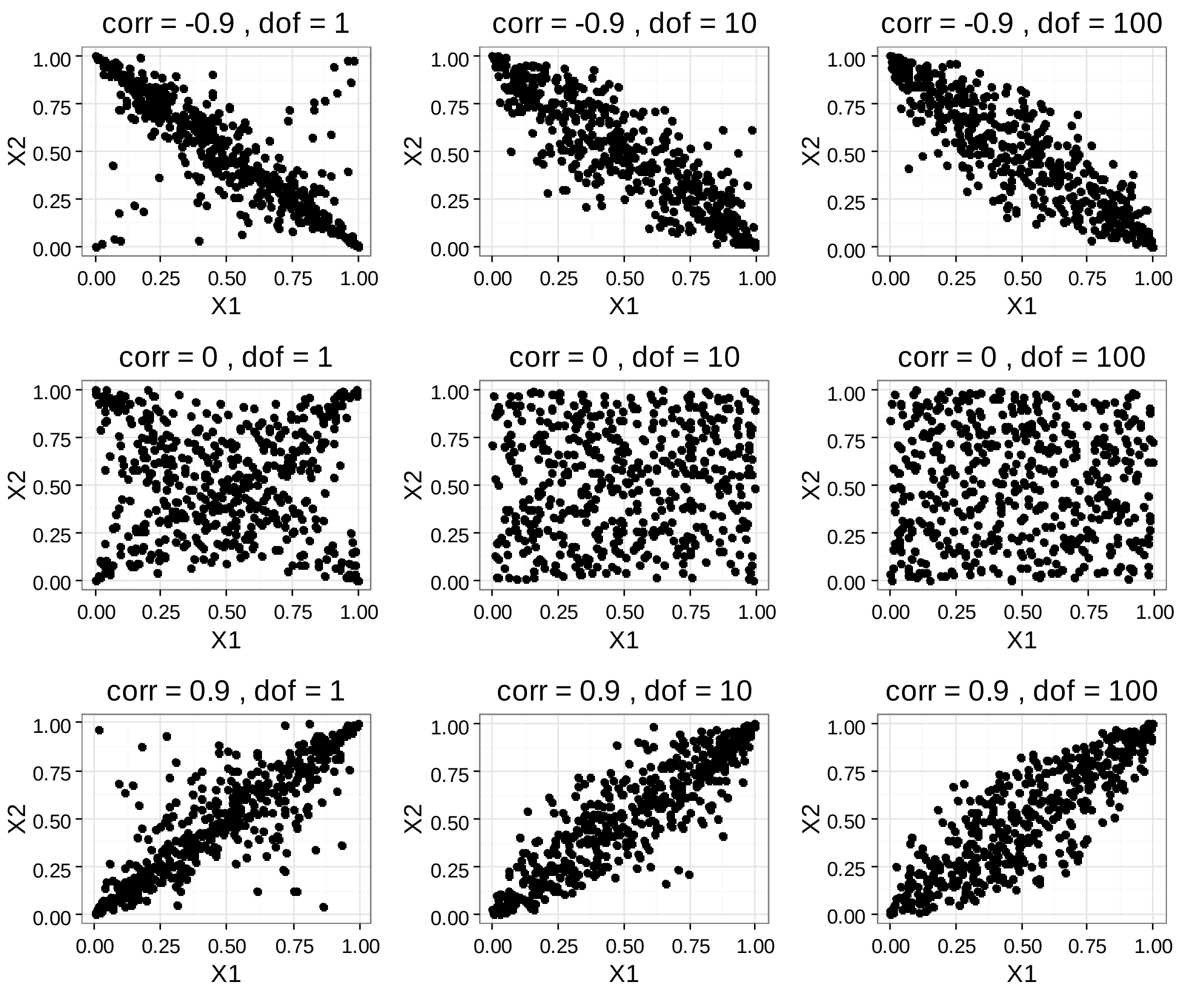
A visual journey of Copulas

Gaussian copula



A visual journey of Copulas

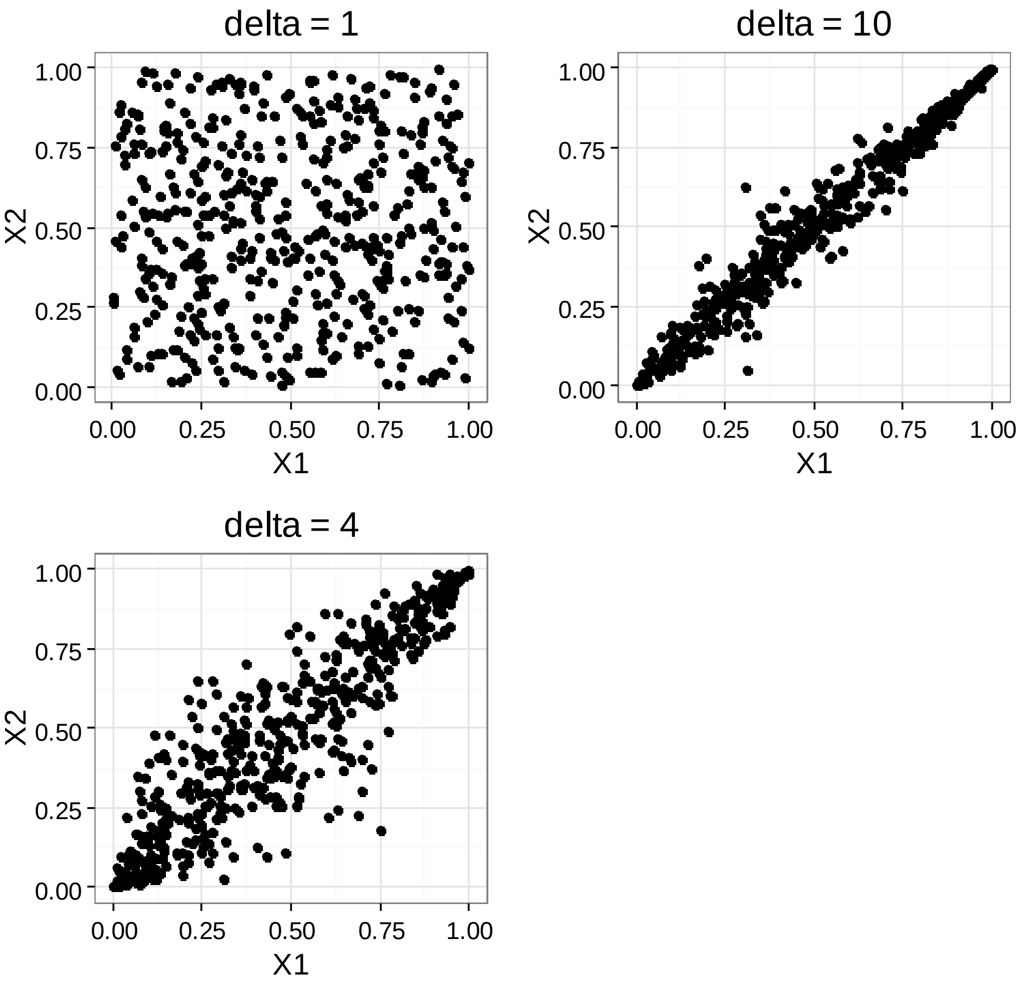
Student-t Copula



UBS

A visual journey of Copulas

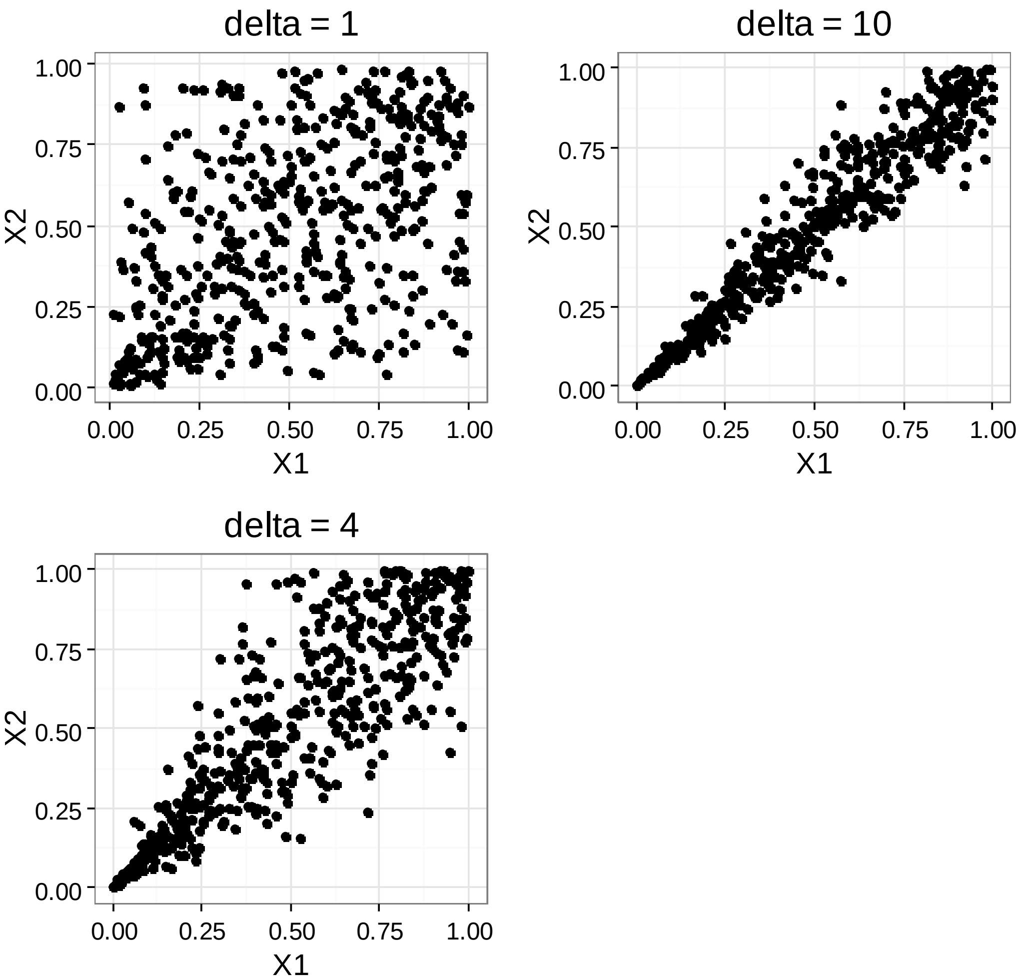
Gumbel copula



UBS

A visual journey of Copulas

Clayton Copula



UBS

Copula simulation

How can we simulate copula?

- Standard approach (known parametric form):
  - Draw from multivariate distribution (multivariate Gaussian, student-t, etc): $F(x_1, \dots, x_k)$;
  - Apply quantile function to each marginal: $F_1^{-1}(x_1), \dots, F_k^{-1}(x_k)$

- CMA (Empirical):
  - Scenarios of joint realisations: $\{x_{1,i}, \dots, x_{k,i}; p_i\}_{i=1,\dots,I}$;
  - Copula is represented by: $\{u_{1,i}, \dots, u_{k,i}; p_i\}_{i=1,\dots,I}$;
  - $u_{n,i} = \sum_{j=1}^{I} p_j \, 1_{\{x_{n,j} \leq x_{n,i}\}}, \forall n = 1, \dots, k, \forall i = 1, \dots, I$

Copula simulation

Step-wise aggregation

- Note we are interested in PnL (loss) distribution: $\mathrm{L} = H(X_1, \dots, X_n)$.
- Let X_{rt} be risk factors for risk type rt: market, credit, operational, etc, $\mathrm{L_{rt}} = H^{rt}(X_{rt})$, L_{rt} is univariate.
- E.g: $F(L_M, L_C) = C_{MC}^{L}\big(F(L_M), F(L_C)\big)$, $L_{MC} = H(L_M, L_C) = L_M + L_C$, L_{MC} is univariate.
- Solving multivariate problem reduced to solving stepwise bivariate problem.

UBS

Thanks for your attention!

Contacts:
alessio.brussino@ubs.com