Lectures in Quantitative Finance

Universität Zürich[UZH]

Modeling Dependencies

Guest Lecture

Pedro Fonseca
Head Business Solutions
Systemorph AG

Zürich, 16 April 2015

Goal

Universität Zürich[UZH]

- This lecture aims at providing a practical overview on the importance of dependencies and how to model them.
- Exercises in Excel (without VBA) will be used to explain the associated techniques. Solutions will be made available.

  If possible, the students should bring a Laptop with Excel.

A *very short list* of software used for calculations:

Excel based:	VBA and @Risk (commercial)
Open source:	R and Octave
Commercial:	Matlab and Mathematica
Library-based:	C#, C++, Java, etc.

Program

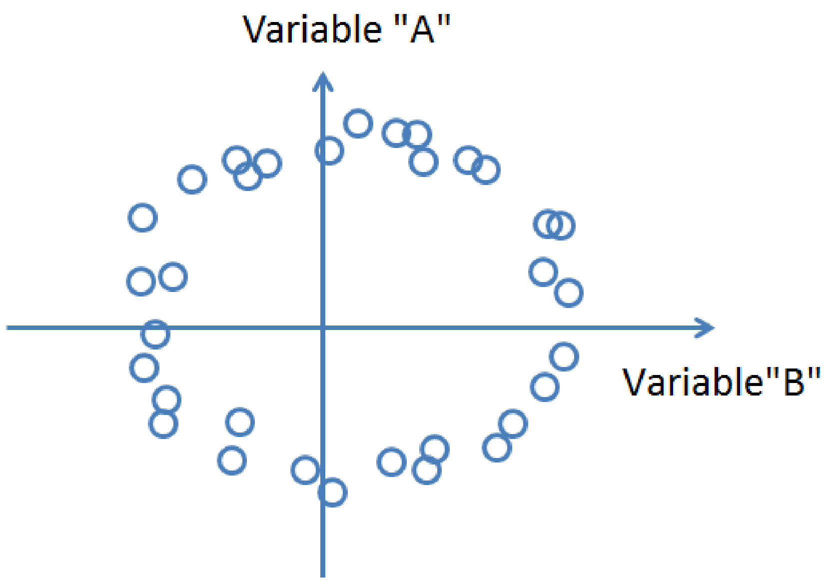
Welcome

1. Introduction to Dependencies
2. Methods for Modeling Dependencies
3. Model Selection and Calibration
4. Trusted Data

Closing Remarks

First, a Notation Issue

- “Correlation” is usually meant as a measure of “dependency”
- “Dependency” is more general
- Indeed, it can happen that dependency exists, but the usual correlation measures are not able to capture it – for instance:



Correlation Measures:

Pearson = zero

Spearman = zero

Kendall = zero

- Dependencies Between What?
- Why Model Dependencies?
- How to Measure Dependencies?

Dependencies Between What?

The short answer (in banking and insurance industries):

"Between risks"

With the usual suspects:

- Credit Risks (counterparty)
- Operational Risks
- Investment Risks (asset)
- Underwriting Risks (liability)

Dependencies Between What?

Universität Zürich[UZH]

The not-so-short answer:

Context
- Event Frequencies
- Severities
- Stock Returns
- …

Range
- Within a Risk
- Across Risks e.g. under-writing, investment, credit, operational

Time Dependency
- Static (Snapshot)
- Across Time

Observation Period
- During 1 Second
- During 5 Minutes
- During 1 Day
- During 1 Year
- …

Time Separation
- At coincident times
- With Time-Lag of 1 Second
- With Time-lag of 1 Year
- …

The main focus during this lecture

Why Model Dependencies?

Example #1 – Mergers

Company Nr. 1

Company Nr. 2

Company Nr. 3

Company Nr. 4

Company Nr. 5

Merged Companies:
New fat-tail risks due to new strong dependencies

Why Model Dependencies?

Dependencies create riskier worlds

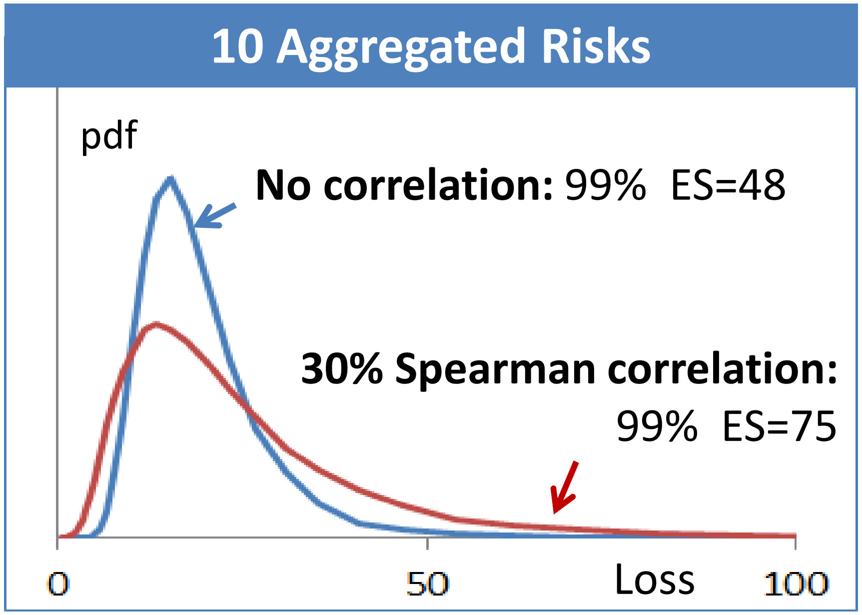
Positive dependency typically **generates fatter-tails**, leading to:

- Less diversification effect
- Higher frequency of “rare” events
- Increased Value-at-Risk (**VaR**)
- Increased Expected-Shortfall* (**ES**)

* Also known as “Tail-Value-at-Risk” or “Conditional-Value-at-Risk”.

Why Model Dependencies?
Example #2 – Less Diversification

If losses from various business lines are dependent,
then the **diversification effect is smaller**.



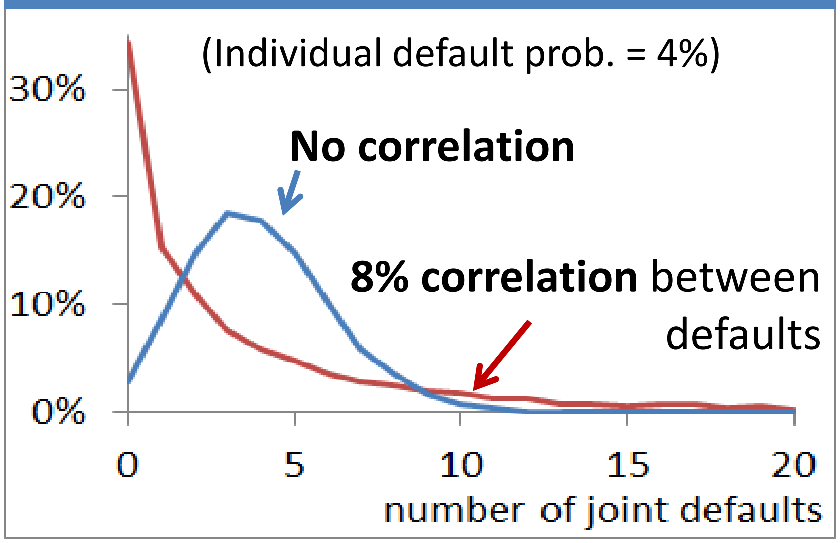
Why Model Dependencies?
Example #3 – Higher Frequency of Rare Events

Universität Zürich[UZH]

If rare events are correlated,

then the **probability of joint events is higher**.

Joint Defaults (100 counterparties)



Why Model Dependencies?

Universität Zürich[UZH]

- Regulation:
  - Basel III and FINMA “Swiss finish” (for Banks)
  - Swiss Solvency Test and Solvency II (for Insurers)
- Reserving / Risk Adjusted Capital
- Pricing
- Capital Allocation
- Business Planning
- Portfolio and Risk Management
- …

To Improve Strategy
(Profitability, Survival, …)

Internal models are important

Why Model Dependencies?

- **Negative dependencies** usually have the opposite effect of positive dependencies, but are less frequent.
- Example: mean-reversal behaviour of stock prices, i.e. relation between the price “before” and “now”.
- (A positive cash-flow means a negative cash-flow for someone else, but that does not count as negative dependency for neither of them.)

Exercise #1

Universität Zürich[UZH]

Random variables can be generated using the fact that the cumulative distribution function $F(x)$ follows a Uniform(0,1) distribution, i.e. $F(x)$~Uniform(0,1).

Generate the variable u from $\mathrm{Uniform}(0{,}1)$ and plug it in the inverse of the cumulative function. The result $x = F^{-1}(u)$ has the desired distribution.

Goals:

1) Generate 100 realizations of a Gaussian variable $X \sim \mathrm{Gaussian}(0, 1)$.
2) Generate 100 realizations of a Poisson variable $\mathrm{Y} \sim \mathrm{Poisson}(3)$.
   Hint: Use a table and the *vlookup* formula to find the inverse $F^{-1}(u)$.
3) Estimate their expected value and standard deviation based on the realizations (the error goes down with larger number of realizations).

How to Measure Dependencies?

Universität Zürich[UZH]

E.g. with **correlation** measures, such as:

Eq. 1

- **Pearson:** (or "linear") $$\rho(X_1,X_2) = \frac{Cov(X_1,X_2)}{\sqrt{Var(X_1)Var(X_2)}}$$

Eq. 2

where $Cov(X_1,X_2) = E(X_1X_2) - E(X_1)E(X_2)$

- Parametric
- Only captures <u>linear</u> dependency

- **Spearman:** $\rho_S(X_1,X_2) = \rho(F_1(X_1), F_2(X_2))$

- **Kendall :** $\tau(X_1,X_2) = E\left[\operatorname{sign}\left((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2)\right)\right]$

where $\tilde{X}_1$ and $\tilde{X}_2$ have the same joint distribution, but are independent of X_1 and X_2.

- Not sensitive to outliers
- Good for heavy-tailed distributions
- Better than linear correlation

Exercise #2

In the **1-Factor Gaussian Model** all variables (e.g. stocks) share a common-factor, the **"market"** M:

$$X_1 = \sqrt{1-\rho}\, Y_1 + \sqrt{\rho}\, M$$
$$X_2 = \sqrt{1-\rho}\, Y_2 + \sqrt{\rho}\, M$$

etc.

All variables have distribution $\mathrm{Gaussian}(0,1)$*. The market M and the idiosyncratic components Y_i are independent.

Goals:

1) Generate 1000 realizations of X_1 and X_2 with $\rho = 30\%$.
2) Estimate the various correlations: a) Pearson (linear); b) Spearman; c) Kendall, and compare with the theory:
   - Pearson: ρ
   - Spearman: $\frac{6}{\pi}\arcsin(\rho/2)$
   - Kendall: $\frac{2}{\pi}\arcsin(\rho)$

* From $\mu + \sigma X_i$ one can get other parameters.

- Explicit vs. Implicit Models
- Common-Factors
- Copulas
- Tail Dependency

Modeling Dependencies

Universität Zürich[UZH]

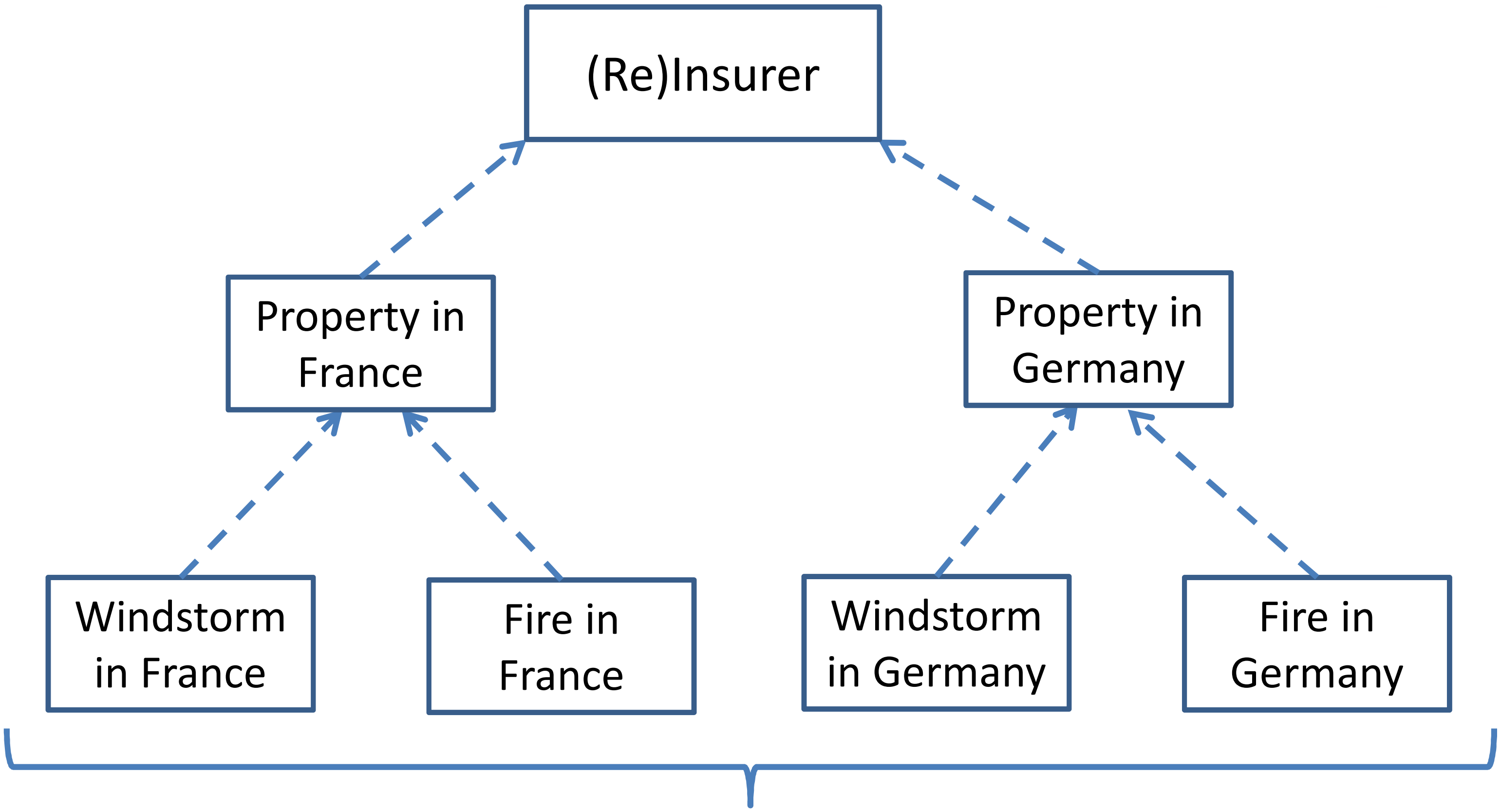
To model independency there is only one choice.

To model dependency there are infinitely many choices.

Explicit vs. Implicit Models

Example #4 – Natural Catastrophes

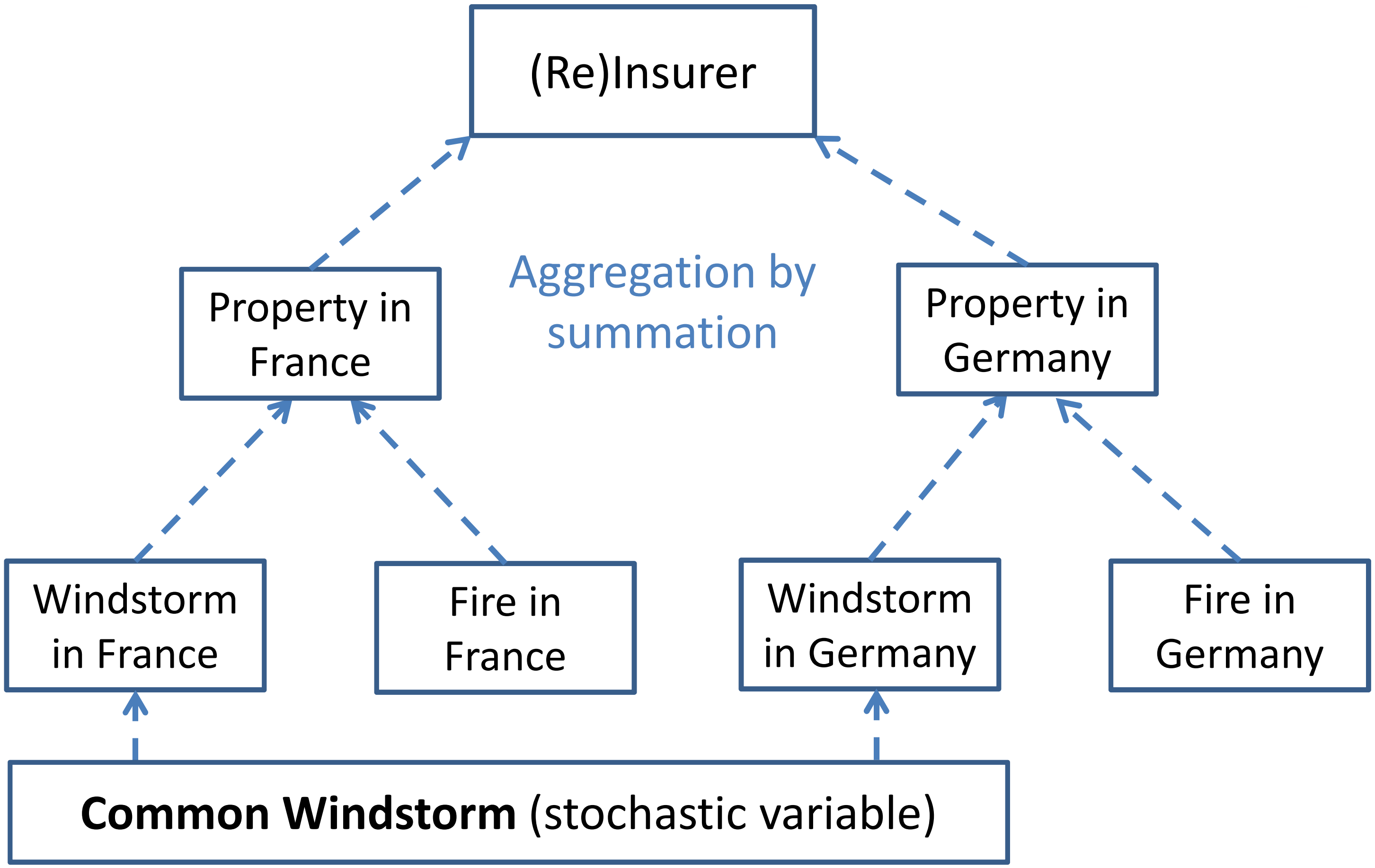
Universität Zürich[UZH]



Not accounting for dependencies can lead to insolvency.

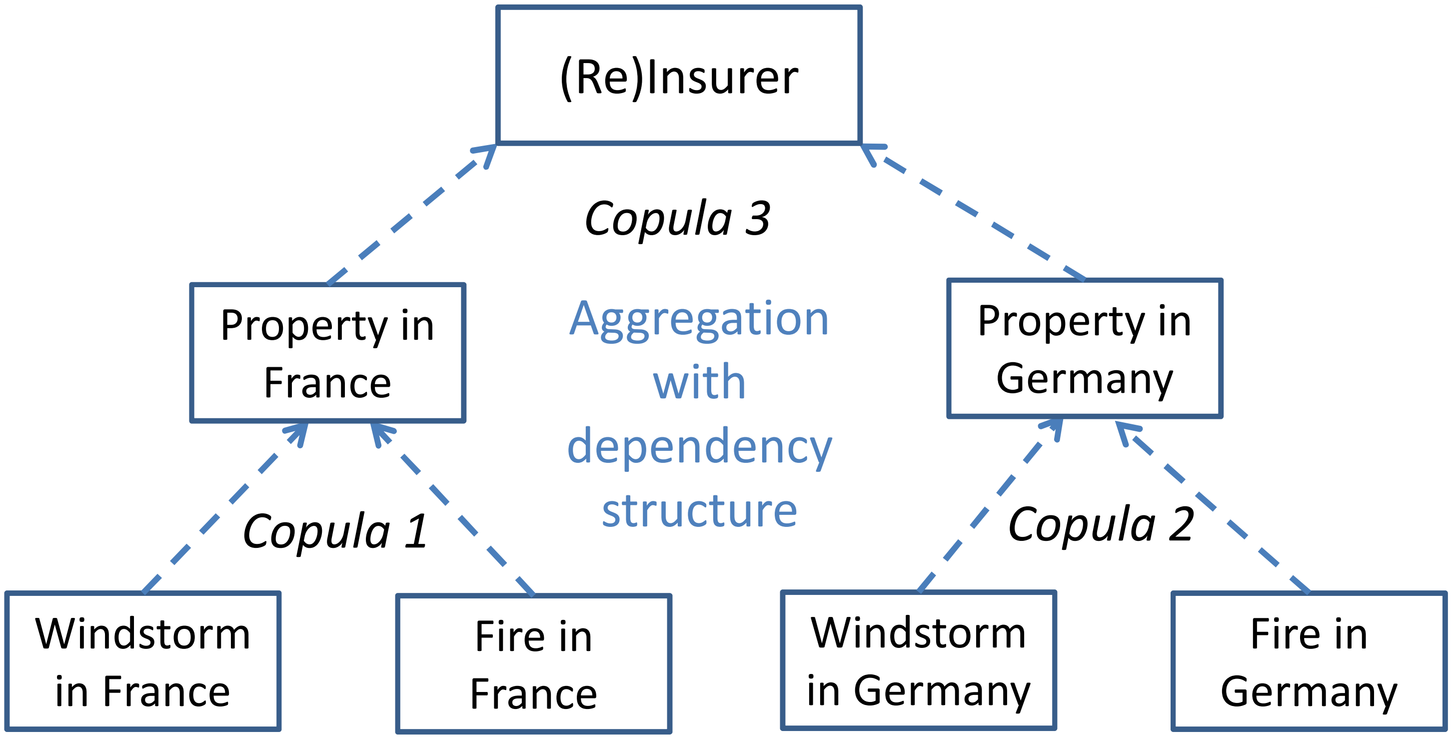
Note: Diagram inspired from work of Roland Bürgi, Michel M. Dacorogna & Roger Iles (2008).

Explicit Models
Example #4 – Common-Factors



Note: Diagram inspired from work of Roland Bürgi, Michel M. Dacorogna & Roger Iles (2008).

Implicit Models
Example #4 – Copulas



Note: Diagram inspired from work of Roland Bürgi, Michel M. Dacorogna & Roger Iles (2008).

Explicit vs. Implicit Models

Universität Zürich[UZH]

Explicit Models		Implicit Models
e.g. Common Factors/Shocks, Causal, …	Regression, Frailty, …	e.g. Copulas, Lévy-Copulas, Pareto-Copulas, …

- Intuitive
- Potentially accurate
- Give insight into business
- But can lead to a false sense of accuracy

- Many types of dependencies
- Explicit tail dependency
- But calibration is complicated,
- and causality might not be known

Implicit Models
Example #5 – Common-Factors

Universität Zürich[UZH]

Same stochastic variable in two or more risks
(e.g. default of a counterparty)

Risks faced by a certain company:

Risk 1: money deposited in Bank A

Risk 2: outstanding receivables from client X, which happens to be the same Bank A

Common-Shock: default of Bank A

Implicit Models
Example #6 – Common-Factors

Adding common stochastic variables to other variables
(e.g. Merton Model, Poisson Shocks), or to their parameters

Risks in a portfolio:

Portfolio: composed by stocks with returns modeled as $X_i \sim \text{Gaussian}$.

1-Factor Gaussian Model: all stocks share a common factor: **the market**.

$$X_1 = \sqrt{1-\rho}\, Y_1 + \sqrt{\rho}\, M$$
$$X_2 = \sqrt{1-\rho}\, Y_2 + \sqrt{\rho}\, M$$

etc.

where the idiosyncratic components $Y_i \sim \text{Gaussian}$ are independent.

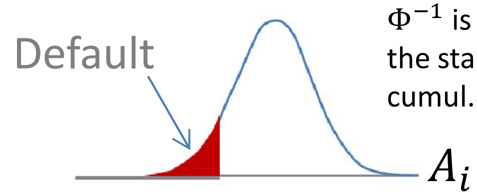
Exercise #3
Merton Model

- The Merton model is commonly used for credit risk (e.g. Basel III)
- It uses the 1-factor Gaussian model (equivalent to the Gaussian copula) to model the assets A_i of the counterparties ($i = 1, 2, \ldots$):

$$A_i = \sqrt{1-\rho}\, Y_i + \sqrt{\rho}\, M$$

  where Y_i are idiosyncratic factors and M the market (see exercise #2).
- Default occurs when the assets A_i drops below a certain value, which is set by the default probability p_i:

$$B_i = \begin{cases} 1 & \text{if} \quad A_i < \Phi^{-1}(p_i) \\ 0 & \text{otherwise} \end{cases} \sim \text{Bernoulli}(p_i)$$



Φ^{-1} is the inverse of the standard normal cumul. distrib. function

Goal: Simulate 3 counterparties with assets correlated with $\rho = 30\%$ and default probabilities $p_i = 5\%$. Estimate numerically the (linear) correlation between defaults.

Hint: Results from exercise #2 can be re-used here.

Note: a similar model was used to produce the chart on page 11.

Exercise #4

Poisson Common-Shocks

Risk 1: Windstorms in France with frequency $X_F \sim \mathrm{Poisson}(5)$, i.e. average of 5 windstorms per year.

Risk 2: Windstorms in Germany with frequency $X_G \sim \mathrm{Poisson}(6)$.

Common-Shock: Pan-European windstorm with frequency $X_C \sim \mathrm{Poisson}(2)$, modeled as:

$$X_F = X_1 + X_C$$
$$X_G = X_2 + X_C$$

Or "**Default**“ instead of "Windstorm" and "**Portfolio**“ instead of "Country"

where $X_1 \sim \mathrm{Poisson}(3)$ and $X_2 \sim \mathrm{Poisson}(4)$.

Goal: Use a Monte-Carlo simulation to estimate the Pearson correlation coefficient. Compare against the theoretical result:

Try to derive this using equations (1-2) in pg. 15

Between 0 and $\min(\lambda_F, \lambda_G)$

$$\rho = \frac{\lambda_C}{\sqrt{\lambda_F \, \lambda_G}}$$

Where λ_F, λ_G and λ_C are the expected values of X_F, X_G and X_C.

Copulas

Universität Zürich[UZH]

- The principle behind copulas is fairly simple. In the bivariate case, consider two random variables X_1 and X_2. A copula C is **a parameterization** of the joint cumulative distribution function:

$$\boxed{F_{1,2}(X_1, X_2) = P(X_1 \leq x_1; X_2 \leq x_2) = \boldsymbol{C}\big(\boldsymbol{F_1}(\boldsymbol{X_1})\,, \boldsymbol{F_2}(\boldsymbol{X_2})\big)}$$

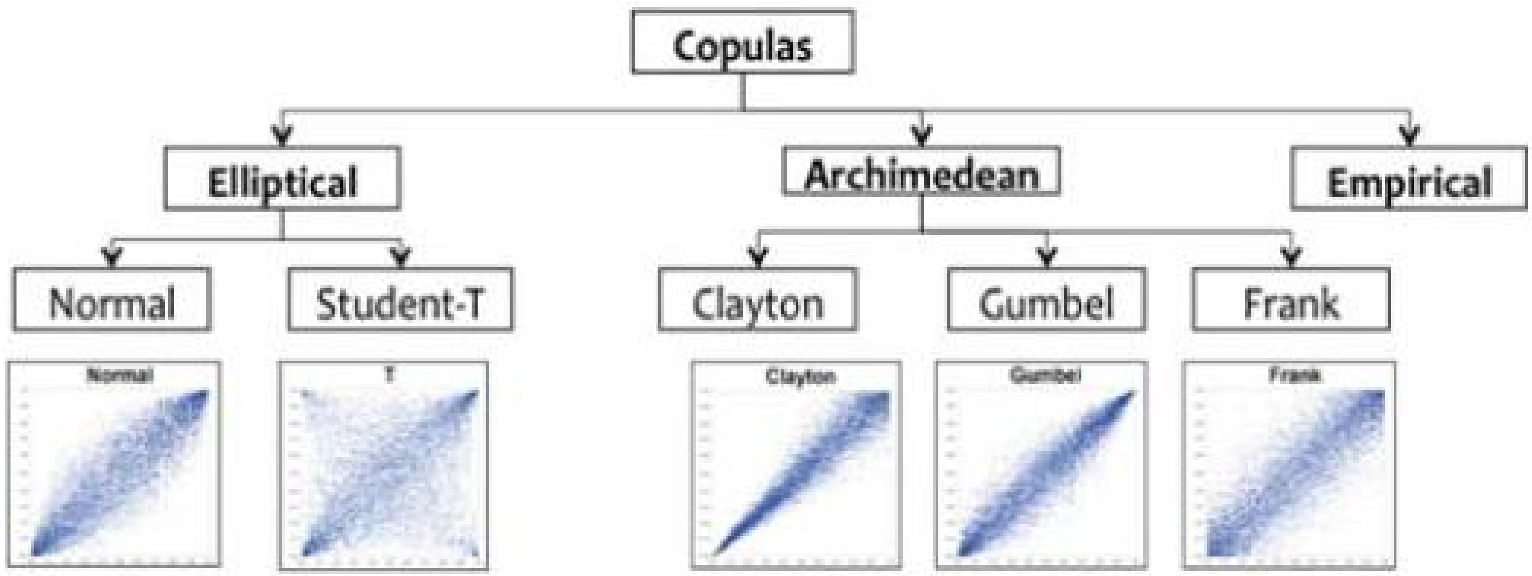
Kendall's tau

- No dependency: $C(u_1, u_2) = u_1 u_2$ like $P(A \wedge B) = P(A)P(B)$
- **FGM***: $C(u_1, u_2) = u_1 u_2 (1 + \theta(1 - u_1)(1 - u_2))$ $|\theta| \leq 1$ and $\tau = \frac{2}{9}\,\theta$
- **Clayton**: $C(u_1, u_2) = \left({u_1}^{-\theta} + {u_2}^{-\theta} - 1\right)^{-1/\theta}$, $\theta \geq 0$ and $\tau = \frac{\theta}{\theta+2}$
- **Gumbel**: $C(u_1, u_2) = \exp\left[-\left((-\ln u_1)^{\theta} + (-\ln u_2)^{\theta}\right)^{1/\theta}\right]$ $\theta \geq 1$ and $\tau = 1 - \frac{1}{\theta}$
- **Gaussian**: $C(u_1, u_2) = \Phi_\rho\big(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\big)$
- **Student's t**: $C(u_1, u_2) = t_{v,\rho}\big({t_v}^{-1}(u_1), {t_v}^{-1}(u_2)\big)$

$|\rho| < 1$ and $\tau = \frac{2}{\pi} \operatorname{asin} \rho$

* Fairly-Gumbel-Morgenstern. Only adequate to model small dependencies.

Copulas
A Short Graphical Overview



Source: SAS

Copulas

Algorithm to generate bivariate copulas*

One uses the fact that derivatives of the copula yield conditional distribution functions – for instance:

$$P\left(X_2 \leq x_2 \middle| X_1 = x_1\right) = \frac{\partial}{\partial u_1} C(u_1, u_2)$$

where $u_1 = F_1(X_1), u_2 = F_2(X_2)$. The outcome follows a Uniform(0,1) and is independent of U_1.

First, generate dependent uniform variables u_1 and u_2:

1. Draw independent v_1 and v_2 from Uniform(0,1).
2. Set $u_1 = v_1$
3. Set $v_2 = \frac{\partial C}{\partial u_1}$ and solve for u_2, i.e. $u_2 = \left(\frac{\partial C}{\partial u_1}\right)^{-1}$, where the right end side contains v_1 and v_2 - see next page for a few examples.

Second, generate the marginal variables x_1 and x_2 :

4. Simply use $x_1 = F_1^{-1}(u_1)$ and $x_2 = F_2^{-1}(u_2)$

* A similar algorithm applies to the multivariate case as well.

Copulas
Algorithm to generate bivariate copulas

Explicit expression for u_2:

- **FGM:** $u_2 = 2v_2/(\sqrt{B} + A)$
  where $A = 1 - \theta(2v_1 - 1)$
  and $B = [1 - \theta(2v_1 - 1)]^2 + 4\theta v_2(2v_1 - 1)$
- **Gumbel:** $\frac{\partial C}{\partial u_1}$ *is not invertible…*
- **Clayton:** $\frac{\partial C}{\partial u_1} = \left[1 + \left(\frac{u_1}{u_2}\right)^{\delta} - u_1{}^{\delta}\right]^{-\frac{1}{\delta}-1}$, so that $v_2 = \frac{\partial C}{\partial u_1}$ yields
  $u_2 = \left[v_1{}^{-\theta}\left(v_2{}^{-\theta/(\theta+1)} - 1\right) + 1\right]^{-1/\theta}$

Exercise #5

Universität Zürich[UZH]

The **losses** from windstorms in France and Germany, S_1 and S_2 respectively, follow a Pareto distribution with scale parameter 3 and shape parameter $\alpha = 4$, i.e.

Or "**Default**“ instead of "Windstorm" and "**Portfolio**“ instead of "Country"

$$F_i(s_i) = 1 - \left(\frac{3}{3+s_i}\right)^4$$

with $s_i > 0$ and $\mathrm{i} = 1{,}2$.

Goal: Introduce a Kendall correlation of 50% between S_1 and S_2 by using the **Clayton copula** with the tail dependency on the upper side, i.e. use the formulas from pages 29-30 and then use

$$\widetilde{u_i} = 1 - u_i$$

to plug in

$$F_i^{-1}(\widetilde{u_i}) = 3(1 - \widetilde{u_i})^{-1/4} - 3$$

Exercise #6

Universität Zürich[UZH]

Optional Slide

The **Gumbel copula** can be generated using the algorithm*:

1. Generate independent $v_1, v_2 \sim$Uniform(0,1).
2. Find w such that $w(1 - \ln(w)\,/\theta) = v_2$, where $0 \le w \le 1$.
3. Set $u_1 = \exp\left[{v_1}^{1/\theta} \ln(w)\right]$ and $u_2 = \exp\left[(1 - v_1)^{1/\theta} \ln(w)\right]$.

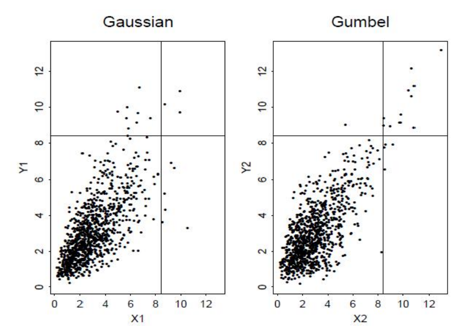
Goal: Take $X_1, X_2 \sim$ Gamma(3,1) and correlate them to 70% (Pearson correlation) using:

a) A Gumbel copula (**Hint:** use a vlookup table to solve for w above);
b) And a Gaussian copula.

For each, estimate:

c) The tail probability $P(X_2 > u \,|\, X_1 > u)$
at $u = F^{-1}(99\%)$.
d) The tVaR of the sum, at 99% confidence level.

Result: while tVaRs are similar, the tail prob. are not.



Source: Embrechts et al, 1998.

* See general Algorithm 6.1. in Embrechts et al (2001) and pages therein for its derivation.

Copulas

Algorithm for Gaussian and Student's t- copulas

Universität Zürich[UZH]

Optional Slide

Given a correlation matrix $\mathbf{\Sigma}$:

Gaussian copula:

1. Perform a Cholesky-decomposition $\mathbf{\Sigma} = \mathbf{A}^T \mathbf{A}$.
2. Generate independent $\tilde{X}_1, \ldots, \tilde{X}_d \sim \text{Gaussian}(0,1)$.
3. Compute $(X_1, \ldots, X_d) = \mathbf{A}\, \widetilde{\mathbf{X}}$.
4. Finally, compute $U_i = \Phi(X_i)$.

Student's t-copula:

Do steps 1 to 3 above.

4. Generate $\xi = \sum_{i=1}^{v} {Y_i}^2$, where $Y_i \sim \text{Gaussian}(0,1)$ are independent.
5. Finally, compute $U_i = t_v(X_i/\sqrt{\xi/v})$, where t_v is the cumulative distribution function of the t-distribution.

Exercise #7

Optional Slide

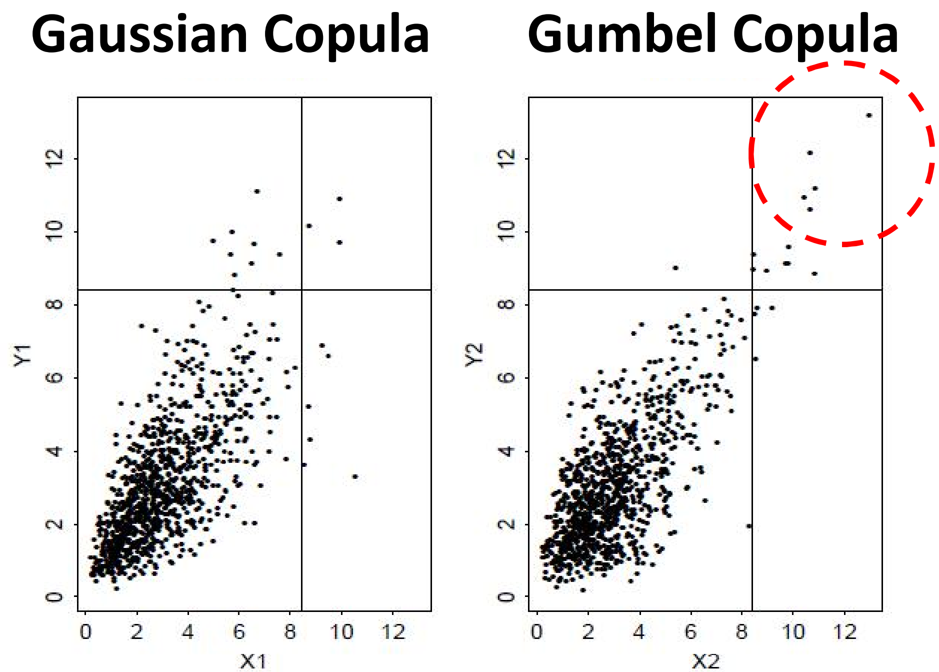
Goal: Simulate the bivariate u_1, u_2 from the Student's t-copula with $\upsilon = 3$ and $\rho = 70\%$.

Compare the estimate of Kendall's tau with the theoretical value.

Tail Dependency

Universität Zürich[UZH]

- Tail dependency can be easily modeled by an appropriate copula.



Source: Embrechts et al, 1998.

- **FGM** copula: None
- **Clayton** copula: On one side
- **Gumbel** copula: On one side
- **Gaussian** copula: None
- **Student's t**-copula: On both sides (non-zero even if $\rho = 0$)

Tail Dependency
Bivariate Case

Optional Slide

- Upper- and Lower-Tail Dependency Coefficients:

$$\lambda_L = \lim_{u \searrow 0} P\left(X_2 \le {F_2}^{-1}(u) \middle| X_1 \le {F_1}^{-1}(u)\right) = \lim_{u \searrow 0} \frac{C(u,u)}{u}$$

$$\lambda_U = \lim_{u \nearrow 1} P\left(X_2 > {F_2}^{-1}(u) \middle| X_1 > {F_1}^{-1}(u)\right) = \lim_{u \nearrow 1} \frac{1 - 2u + C(u,u)}{1-u}$$

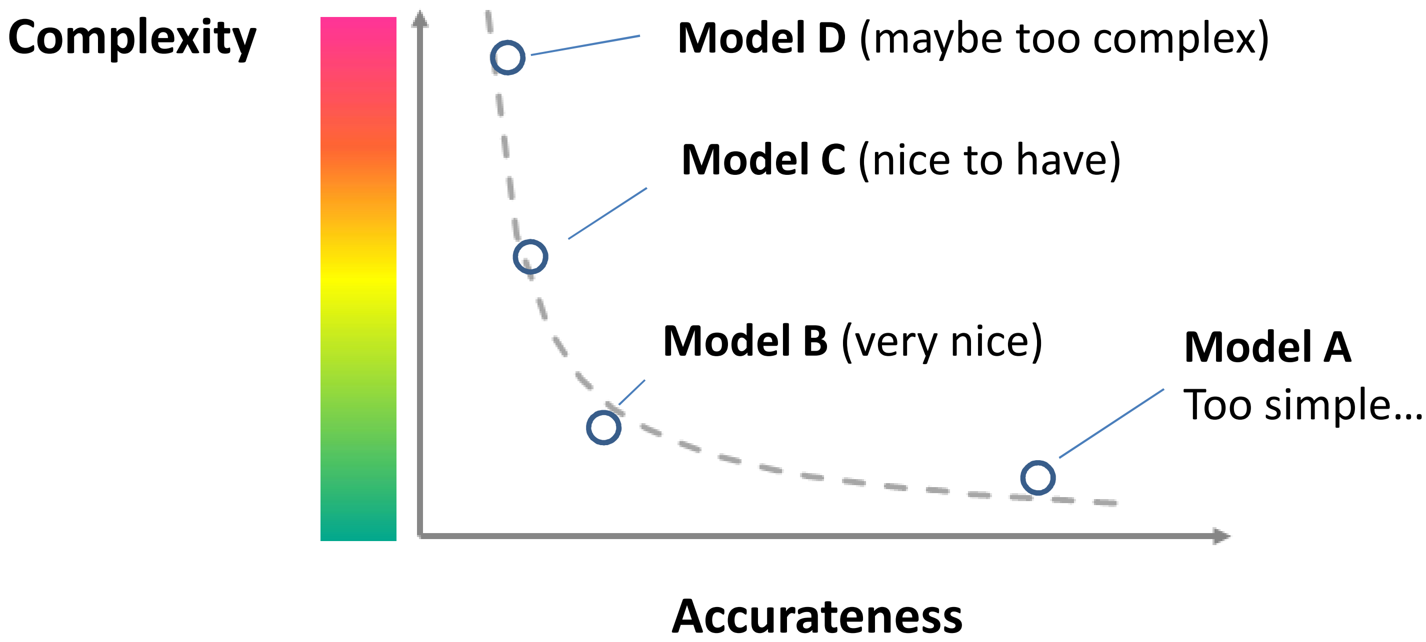
- **FGM** copula: $\lambda_L = 0$ and $\lambda_U = 0.$
- **Clayton** copula: $\lambda_L = 2^{-1/\theta}$ and $\lambda_U = 0.$
- **Gumbel** copula: $\lambda_L = 0$ and $\lambda_U = 2 - 2^{1/\theta}.$
- **Gaussian** copula: $\lambda_L = 0$ and $\lambda_U = 0.$
- **Student's t**-copula: $\lambda_L = \lambda_U = 2t_{\upsilon+1}\left(-\sqrt{\frac{(\upsilon+1)(1-\rho)}{1+\rho}}\right).$ Non-zero even if $\rho = 0$

- Model Selection
- Model Calibration

Model Selection

Universität Zürich[UZH]

- A model should be a good balance between accurateness and complexity.



- Can the model be simplified?
- Can business and management understand the model?
- Was the model selection properly documented?

Model Selection

Universität Zürich[UZH]

Optional Slide

- Modeling with **common-factors** provides with a relatively natural selection of the model and a path to sequential improvements.
- When modeling with **copulas**, the large number of choices can actually be a disadvantage.

A possible procedure for selecting an appropriate copula:

- Start by choosing/fit the marginal distributions.
- Then choosing a copula which will bring the desired/expected dependency structure, e.g. based on tail dependency. There are also various statistical tests: Akaike information criterion, Pseudo-likelihood ratio tests, Bayes factor.

Model Calibration

Universität Zürich[UZH]

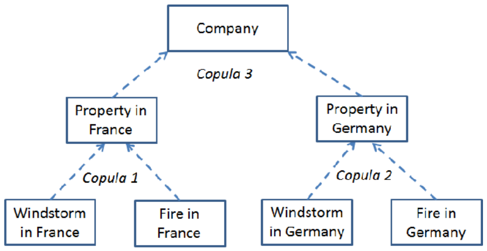
- One has to pay close attention to the fact that dependencies:
  - Are subject to **measure uncertainties**
  - Usually **change over time** (e.g. increasing in times of stress)
  - Might result from **spurious relationships**
- Visual inspection is important
- Calibration of **common-factors** parameters is usually based on exposure/volume data, market data and expert estimates.
- Calibration of **copula** parameters typically uses:
  - Method of moments. For instance, estimating Kendall's tau
  - Maximum likelihood estimation

Model Calibration
Dependency Structure

Optional Slide

- It is difficult to calibrate a **correlation matrix**:
  - It usually has many entries

    E.g. given N risks or risks elements, there are $N(N-1)/2$ pair-wise correlations (e.g. 435 entries for 30 risks).
  - It *must* be symmetric and positive-semidefinite, i.e. cannot have negative eigenvalues.
  - The entries will have a lot of random error



- Trusted Data
- Overview of Systemorph

Why is trusted data important?

Universität Zürich[UZH]

- Good decisions and good modelling need comprehensive and reliable data
- Growing data sources, competition and reporting requirements – e.g.:

SST	Swiss Solvency Test (2006)
Solvency I / II	EU Directive (1973 / 2016)
ORSA	Own Risk and Solvency Assessment (2015)
BCBS 239	Effective risk data aggregation and risk reporting (2016)
Basel III	Third Basel accord (2019), plus Swiss finish
NBA / NBO	(Swiss) National Bank Act / Ordinance (2004, 2013)
MiFID I / II	Markets in Financial Instruments Directive (2007 / ?)
Dodd-Frank	US regulation on OTCs (2010 & ongoing)
EMIR I / II	European regulation on OTCs (2012 & ongoing)
FATCA	Foreign Account Tax Compliance Act (2014 & ongoing)
IFRS4	Intl. Financial Reporting Standard 4 (2018?)
etc.	

Potential game changers

Systemorph at a Glance

SYSTEMORPH®

Background

- Founded 2011, headquartered in Zurich

Customers

- Global financial services firms, focus on (re)insurance and banks

Mission

- Revolutionize software solutions for financial institutions
- Streamline and simplify risk reporting, analysis and actuarial functions

Team

- Risk and capital management, modeling, enterprise systems, information management
- All hold advanced degrees in computer science, physics or mathematics

Do you trust your data?

SYSTEMORPH®

Who owns them?

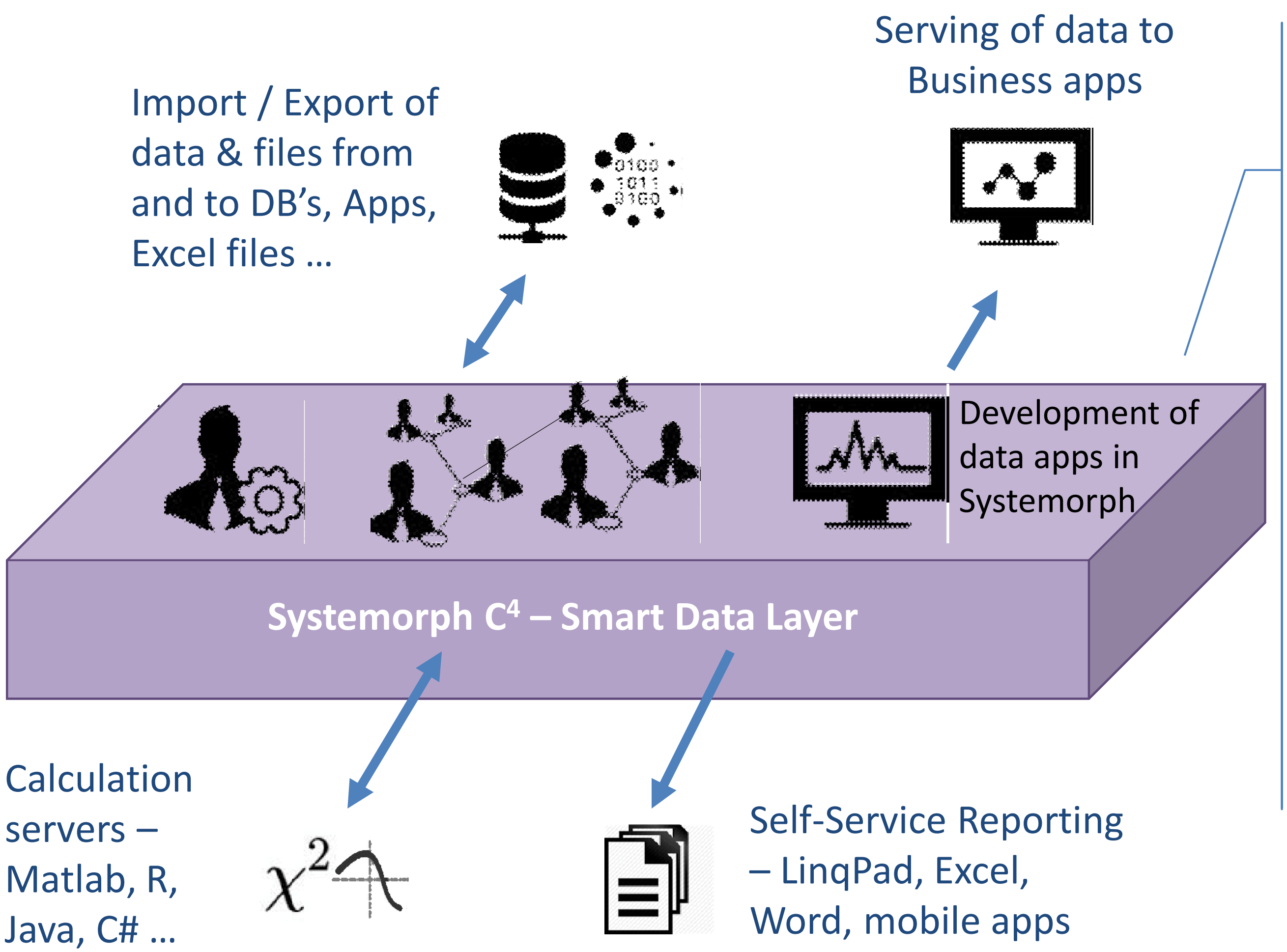
How are they changing?

Are they accurate?

How do we use them for modelling?

How do we share them?

C^4 – Data Intelligence Re-imagined …



- Link data assets
- Build powerful apps
- Decentralize ownership
- Collaborate
- Integrate analytics
- Track data history and changes
- Manage data quality
- Report to stakeholders

Trusted data!

- Closing Remarks
- Literature

Closing Remarks
Principles for effective risk aggregation and risk reporting

Universität Zürich[UZH]

- Following the 2007 financial crisis, the Bank for International Settlements has issued Basel III, and **BCBS 239** (Principles for effective risk data aggregation and risk reporting):
- *"One of the most significant lessons learned from the [2007] global financial crisis [was that] many banks lacked the ability to aggregate risk exposures and identify concentrations quickly and accurately at the bank group level, across business lines and between legal entities.*

  Some banks were unable to manage their risks properly because of weak risk data aggregation capabilities and risk reporting practices. This had severe consequences to the banks themselves and to the stability of the financial system as a whole."
- **Systemically important** banks will comply first. Others will follow.

Literature
A Very Short List

General:

- "Measurement and Modelling of Dependencies in Economic Capital", R.A. Shaw, A.D. Smith and G.S. Spivak, British Actuarial Journal**,** 2010.
- "Credit Portfolio Modeling Handbook", R. Martin, Credit Suisse, 2004.

Copulas:

- "Coping with Copulas", T. Schmidt, 2006.
- "Copula Modeling: An Introduction for Practitioners", P.K. Trivedi and D. M. Zimmer, 2005.
- "Modelling Dependence with Copulas and Applications to Risk Management", P. Embrechts, F. Lindskog and A. McNeil, 2001.
- "Correlations and Dependency in Risk Management: Properties and Pitfalls", P. Embrechts, A. McNeil and D. Straumann, 1998.

Universität Zürich[UZH]

Thank you very much for your attention.

If you have questions, comments or suggestions, please contact:

Pedro Fonseca
Head Business Solutions
Systemorph AG
pfonseca@systemorph.com