Institute and Faculty of Actuaries

Mission Impossible 5: Resolving the Copula Paradox

Stephen Makin
David Stevenson

Birmingham
11 November 2014

Agenda

1. Introduction and background
2. Theory vs. practice
3. The choices available
4. Fitting and validating
5. Towards a conclusion – the role of judgement
6. Next steps for our work

Appendix: copulas – technical background material
Appendix: textbook example
Appendix: maximum likelihood

2

Institute and Faculty of Actuaries

1. Introduction and background

Copulas – a reminder

What is a copula?
- A d-dimensional copula is a multivariate distribution function on $[0,1]^d$ with uniform marginals
- Example: two-dimensional independence copula $C(u,v) = uv$

Sklar's theorem – key result

F a joint distribution function with continuous marginals $F_1, \dots, F_d$:

$$C(u_1, \dots, u_d) = F\left(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\right)$$

Tail dependence
- Limiting conditional probability of joint extreme events
- Coefficients of tail dependence zero for Gaussian, non-zero for Student T

Why copulas are useful?
- Application in internal models
- SCR = 99.5th percentile of $L(X_1, \dots, X_d)$
- Copulas provide means of separating the loss function L from the risk factor distribution $(X_1, \dots, X_d)$ and applying simulation techniques to generate a full PDF
- L can be estimated using a proxy model
- Models for the distribution functions of $X_1, \dots, X_d$ and an algorithm for simulating values $(u_1, \dots, u_d)$ from a copula, provide a recipe for generating simulations of the $(X_1, \dots, X_d)$:

$$(u_1, \dots, u_d) \to (X_1, \dots, X_d) = \left(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\right)$$

Institute and Faculty of Actuaries

2. Theory vs. practice

Theory vs. practice

Textbook

- Low dimensions
- Dependency between similar risks (e.g. between equity stocks, or between exchange rates), often constituents of an index
- Large volume of data
- Consistent data periods and frequencies (e.g. daily)
- Homogeneity assists model selection and fitting and guarantees coherence (e.g. PSD correlation matrices)

In practice

- Higher dimensions
- More disparate risks (e.g. equity returns vs. interest rates, bond spreads vs. persistency)
- Limited data (or, sometimes, no useful data at all)
- Inconsistent time periods $\Rightarrow$ need to use time intervals where data overlaps
- Sometimes inconsistent frequencies
- Parameter values vary significantly by time period
- Have to parameterise model "bit by bit" (e.g. correlations estimated over different data periods) or using judgement to compensate for lack of data
- Can lead to inconsistencies that require adjustments (e.g. non-PSD correlation matrices)

6

Institute and Faculty of Actuaries

3. The choices available

Implicit copulas

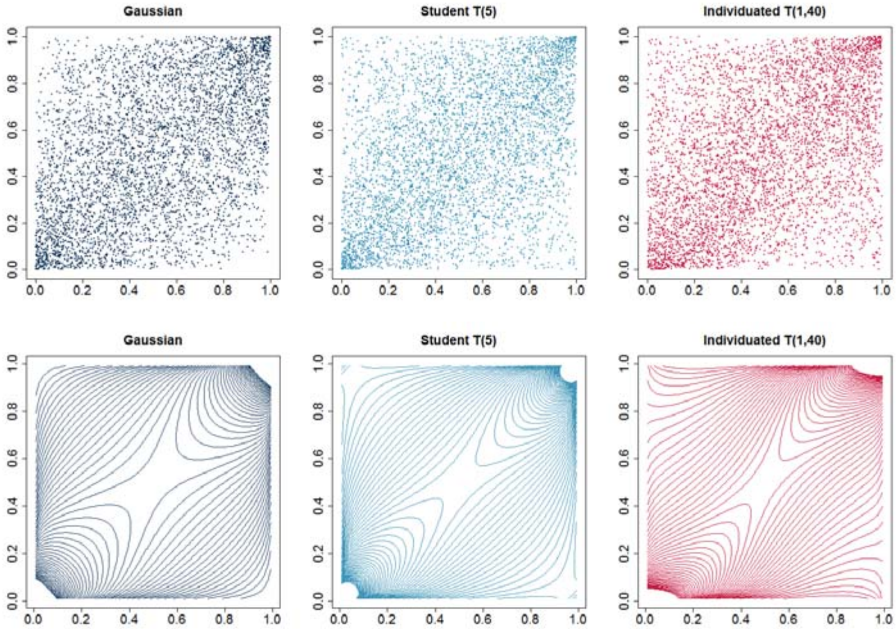
- Copulas underlying well-known multivariate distributions, e.g. Gaussian or Student T
- Defined by $C(u_1, \dots, u_d) = F\left(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\right)$
- Straightforward to simulate from.

- T copula:
  - Simulate Z from $N(0, \Sigma)$ (e.g. using Cholesky) and independently U from $U[0,1]$.
  - Set $W = G_\nu^{-1}(U)$ where G_ν is the distribution function of χ_ν^2
  - Set $(u_1, \dots, u_d) = \left(t_\nu^{-1}\left(\sqrt{\frac{\nu}{W}}z_1\right), \dots, t_\nu^{-1}\left(\sqrt{\frac{\nu}{W}}z_d\right)\right)$

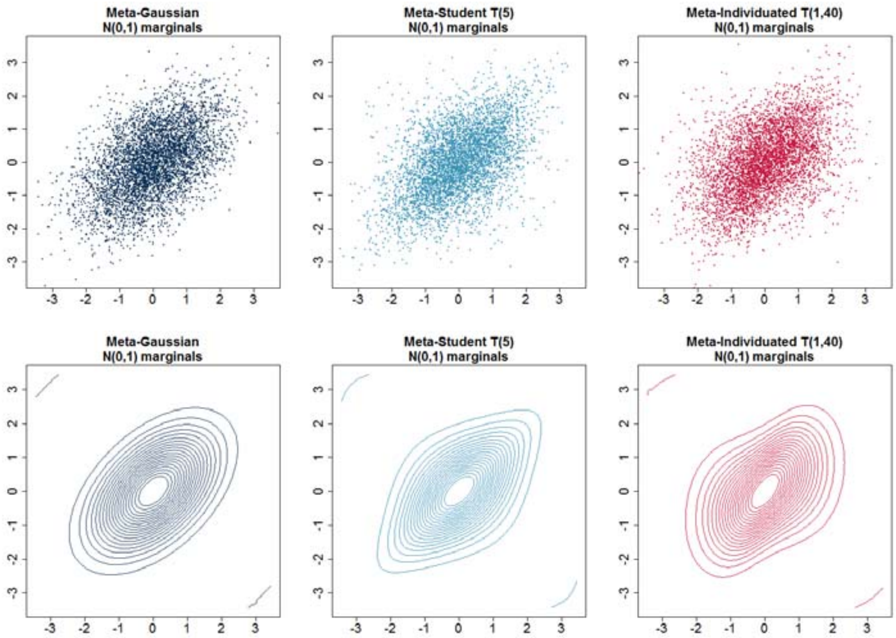
- Individuated T (generalisation of T)
  - Set $W_i = G_{\nu_i}^{-1}(U)$ where G_{ν_i} is the distribution function of $\chi_{\nu_i}^2$
  - Set $(u_1, \dots, u_d) = \left(t_{\nu_1}^{-1}\left(\sqrt{\frac{\nu_1}{W_1}}z_1\right), \dots, t_{\nu_d}^{-1}\left(\sqrt{\frac{\nu_d}{W_d}}z_d\right)\right)$

- Large number of free parameters (e.g. correlation matrix Σ and degrees of freedom)
- Exhibit symmetry due to presence of quadratic form (elliptical – N and T, radial – IT)
- Student T and Individuated T exhibit non-zero tail dependence

8


Gaussian
Student T(5)
Individuated T(1,40)
Gaussian
Student T(5)
Individuated T(1,40)

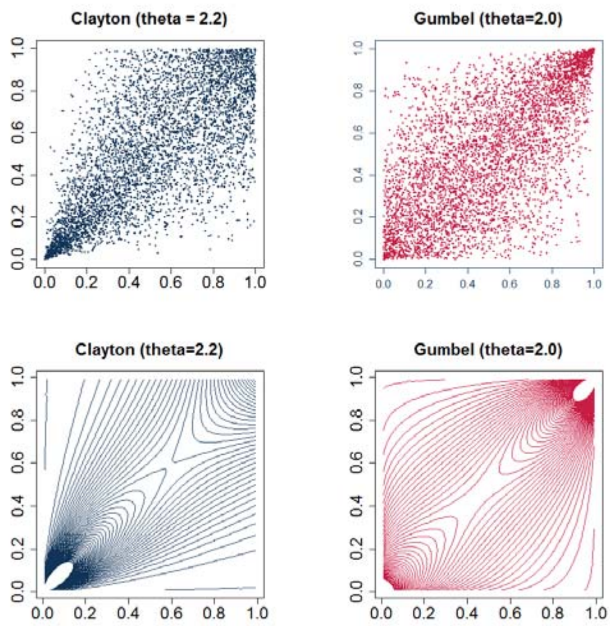
9

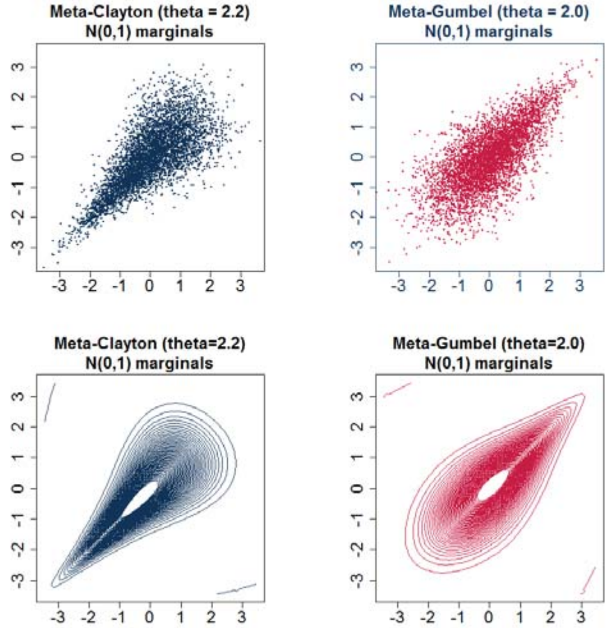

Meta-Gaussian
N(0,1) marginals
Meta-Student T(5)
N(0,1) marginals
Meta-Individuated T(1,40)
N(0,1) marginals
Meta-Gaussian
N(0,1) marginals
Meta-Student T(5)
N(0,1) marginals
Meta-Individuated T(1,40)
N(0,1) marginals

10

Explicit copulas

- Copula defined directly by a function
- Example – Archimedean copula
- $C_\varphi(u_1, \dots, u_d) = \varphi^{-1}\big(\varphi(u_1) + \cdots + \varphi(u_d)\big)$ where $\varphi: [0,1] \to [0, \infty)$ is a strictly decreasing function (the generator function) with $\varphi(0) = \infty$ and $\varphi(1) = 0$
- Large family, including Gumbel, Clayton, Frank, FGM, etc
  - Gumbel $C_\theta(u_1, u_2) = \exp(-\left[(-\log u_1)^\theta + (-\log u_2)^\theta\right]^{1/\theta}$
  - Clayton $C_\theta(u_1, u_2) = \left[u_1^{-\theta} + u_2^{-\theta} - 1\right]^{-1/\theta}$
- More challenging to simulate from
- Limited number of free parameters $\Rightarrow$ less flexibility in modelling dependency in higher dimensions, but exhibit tail dependence
- Strong symmetry (under permutation of variables)





Empirical copulas

- Step functions defined by data
- “Lumpy” when historical data is used
- Outputs therefore sensitive to data updates
- Can replace historical data with synthetic data generated using an ESG
- ESG output can be augmented to include insurance and other risks in addition to economic risks
- Could use Iman-Conover method as way of gluing ranks of simulated marginal distributions in line with ranks of simulated pseudo-observations from copula (see e.g. Mildenhall)
- Not considered here but could be included in further work

Choosing a copula
In practice

- Archimedean copulas less useful for applications in life assurance as they lack the number of free parameters required to reproduce range of correlations required in high dimensional problems
- May be more relevant for some non-life applications
- Dimensionality and modelling practicality reduces choice to elliptic family and generalisations – Gaussian, T and IT in practice
- Balance additional complexity with increased subjectivity of parameterisation against background of limited data
- Consider Use Test: transparency and communication
- Compensate for limitations (e.g. using Gaussian with stronger correlations)
- Vast majority of firms have opted for Gaussian, very few for T or IT

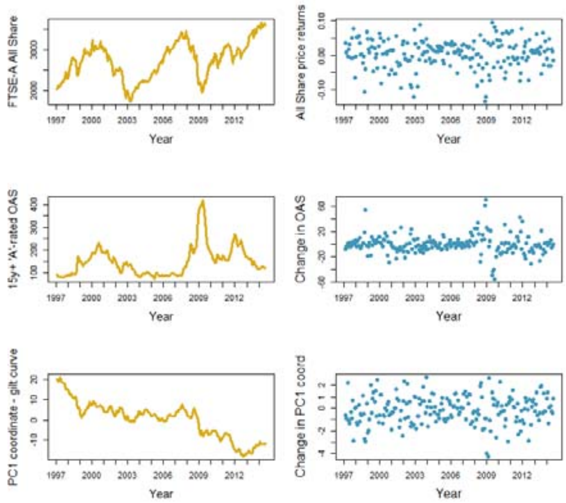
15

Institute and Faculty of Actuaries

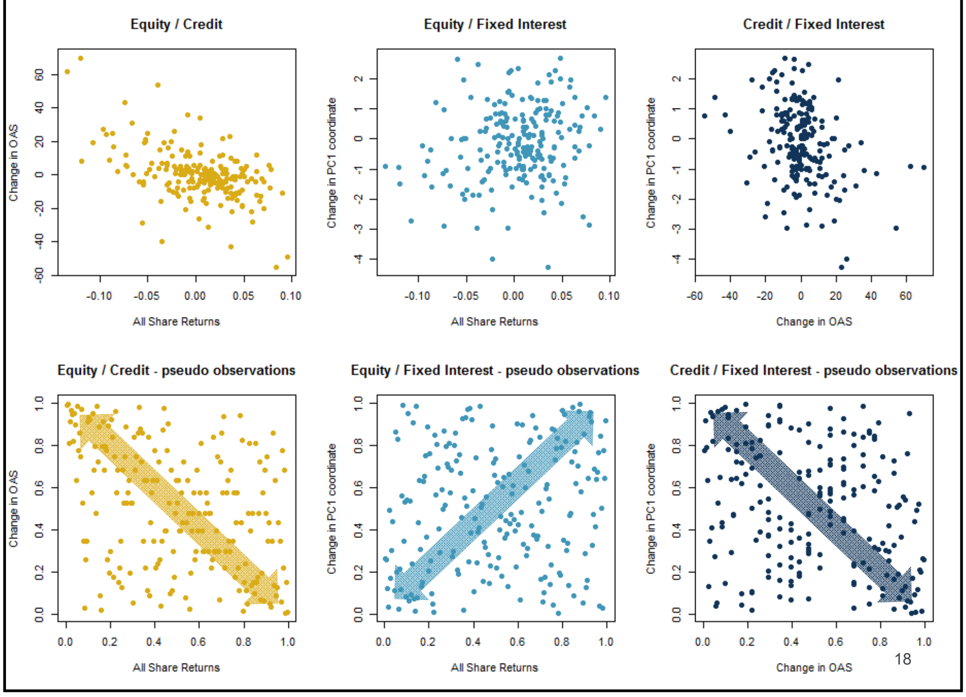
4. Fitting and validating

A case study: equity, credit, fixed interest

- 211 data points: monthly from 31 December 1996 to 30 June 2014
- 210 returns: monthly from January 1997 to June 2014
- Equity: returns on FTSE-A All Share price index
- Credit: changes in OAS on 15 year+ ‘A’-rated credit
- Fixed interest: changes in PC1 coordinate on nominal gilt curve



17



18

Approach

- Focus on Gaussian, T and Individuated T copulas

Copula	Correlation	Degrees of freedom
Gaussian	Yes – d×d matrix	n/a
Student T	Yes – d×d matrix	Yes – one
Individuated T	Yes – d×d matrix	Yes – one for each risk

Test for evidence of dependence → Estimate parameters (maximum likelihood) → Review fitted copulas

19

Test for evidence of dependence

- Null hypothesis H_0: independence copula applies
- Under H_0, Spearman ~ $N(0, 1/\sqrt{n-1})$ as $n \to \infty$
- n=210 $\Rightarrow$ can't reject H_0 if Spearman in [-0.136, 0.136]

Risk Pair	Spearman	95% CI	P-value
Equity / Credit	-0.426	[-0.562, -0.290]	7.3×10^{-10}
Equity / PC1	0.165	[0.029, 0.300]	0.017
Credit / PC1	-0.275	[-0.410, -0.139]	7.1×10^{-5}

- Strong evidence of non-trivial dependence
- Other tests based on other correlation measures show similar results

20

Parameter estimation
Individuated T density

$$c_{\nu}^{\Sigma}(\boldsymbol{u}) = \left[\int_0^1 \frac{\varphi_\Sigma\left(\frac{t_{\nu_1}^{-1}(u_1)}{G_{\nu_1}^{-1}(s)}, \dots, \frac{t_{\nu_d}^{-1}(u_d)}{G_{\nu_d}^{-1}(s)}\right)}{\prod_{j=1}^d G_{\nu_j}^{-1}(s)} ds\right] \Bigg/ \left[\prod_{j=1}^d \left(1 + \frac{t_{\nu_j}^{-1}(u_j)^2}{\nu_j}\right)^{-\frac{(\nu_j+1)}{2}} \frac{\Gamma\left(\frac{\nu_j+1}{2}\right)}{\Gamma\left(\frac{\nu_j}{2}\right)\sqrt{\nu_j \pi}}\right]$$

where

- $\varphi_\Sigma(\boldsymbol{z}) = \exp\left(-\frac{1}{2}\boldsymbol{z}^T \Sigma^{-1} \boldsymbol{z}\right) / \sqrt{(2\pi)^d \det(\Sigma)}$ is the multivariate normal density
- $t_\nu^{-1}(u)$ is the inverse of the standard Student's T density with ν degrees of freedom
- $G_\nu^{-1}(s)$ is the inverse distribution of $\sqrt{\nu / \chi_\nu^2(s)}$

21

Parameter estimation
Individuated T log likelihood

$$\sum_{i=1}^n \log\left[\int_0^1 \frac{\varphi_\Sigma\left(\frac{t_{\nu_1}^{-1}(u_{i,1})}{G_{\nu_1}^{-1}(s)}, \dots, \frac{t_{\nu_d}^{-1}(u_{i,d})}{G_{\nu_d}^{-1}(s)}\right)}{\prod_{j=1}^d G_{\nu_j}^{-1}(s)} ds\right]$$
$$+ \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^d (\nu_j + 1) \log\left(1 + \frac{t_{\nu_j}^{-1}(u_{i,j})^2}{\nu_j}\right)$$
$$- n\sum_{j=1}^d \log\Gamma\left(\frac{\nu_j + 1}{2}\right) + n\sum_{j=1}^d \log\Gamma\left(\frac{\nu_j}{2}\right) + \frac{1}{2}n\sum_{j=1}^d \log \nu_j$$
$$+ \frac{1}{2} nd \log \pi$$

22

Results
Numerical techniques applied

- Gaussian and Standard T: standard libraries and functions in R
- Individuated T: we had to write these ourselves

	Eq / Cr			Eq / PC1			Cr / PC1		
	Rho	Nu1	Nu2	Rho	Nu1	Nu2	Rho	Nu1	Nu2
Gaussian	-0.489	-	-	0.180	-	-	-0.317	-	-
T	-0.463	2.589	-	0.176	6.046	-	-0.308	8.748	-
IT	-0.464	2.465	2.690	0.218	39.30	0.577	-0.307	5.641	11.670

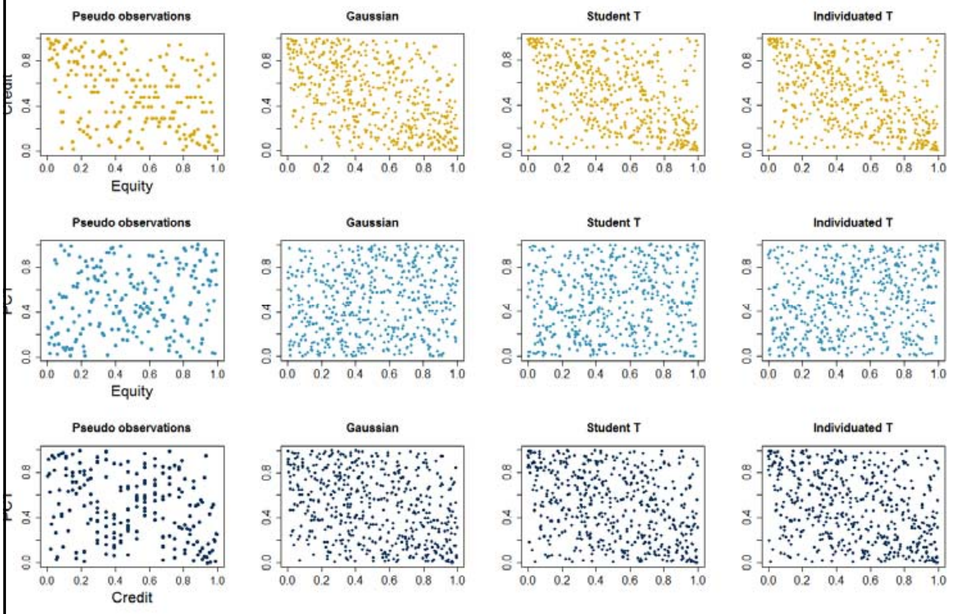
Eq / Cr / PC1	Rho	Nu1	Nu2	Nu3
Gaussian	$\begin{bmatrix} 1 & & \\ -0.489 & 1 & \\ 0.183 & -0.318 & 1 \end{bmatrix}$	-	-	-
T	$\begin{bmatrix} 1 & & \\ -0.484 & 1 & \\ 0.159 & -0.281 & 1 \end{bmatrix}$	4.440	-	-
IT	$\begin{bmatrix} 1 & & \\ -0.463 & 1 & \\ 0.151 & -0.288 & 1 \end{bmatrix}$	2.510	2.885	11.708

Review fitted copulas
Filtering techniques

Risk Pair	Gaussian	Standard T	Individuated T
Equity / Credit			
Equity / PC1			
Credit / PC1			
Eq / Cr / PC1			

Is the increasing complexity justified?

- Visually and numerically
  - AIC = -2logL + 2p where L = likelihood, p = number of parameters
  - minimise AIC: Model A preferred to Model B if $AIC_A < AIC_B$
- Model A has p_A parameters, Model B has p_B ($>p_A$) parameters, A is a special case of B
  - D = 2(L_B – L_A) ~ $\chi^2_{(p_B - p_A)}$
  - 95th percentile of χ^2_1 is 3.84 and 95th percentile of χ^2_2 is 5.99.



Review fitted copulas

Filtering – bivariate

Eq / Cr	Rho	Nu1	Nu2	log L	AIC	D (>3.84)
Gaussian	-0.489	-	-	27.35	-52.70	-
T	-0.463	2.589	-	34.05	-64.10	13.40
IT	-0.464	2.465	2.690	34.06	-62.11	0.01

Eq / PC1	Rho	Nu1	Nu2	log L	AIC	D (>3.84)
Gaussian	0.180	-	-	3.36	-4.72	-
T	0.176	6.046	-	4.71	-5.42	2.71
IT	0.218	39.30	0.577	5.61	-5.23	1.81

Cr / PC1	Rho	Nu1	Nu2	log L	AIC	D (>3.84)
Gaussian	-0.317	-	-	10.42	-18.84	-
T	-0.308	8.748	-	10.97	-17.94	1.10
IT	-0.307	5.641	11.670	10.99	-15.98	0.04

Review fitted copula
Filtering – trivariate

Copula	Rho	Nu1	Nu2	Nu3	log L	AIC	D (>6)
Gaussian	$\begin{bmatrix} 1 & & \\ -0.489 & 1 & \\ 0.183 & -0.318 & 1 \end{bmatrix}$	-	-	-	38.02	-70.03	-
T	$\begin{bmatrix} 1 & & \\ -0.484 & 1 & \\ 0.159 & -0.281 & 1 \end{bmatrix}$	4.440	-	-	44.21	-80.42	12.39
IT	$\begin{bmatrix} 1 & & \\ -0.463 & 1 & \\ 0.151 & -0.288 & 1 \end{bmatrix}$	2.510	2.885	11.708	45.79	-79.58	3.16

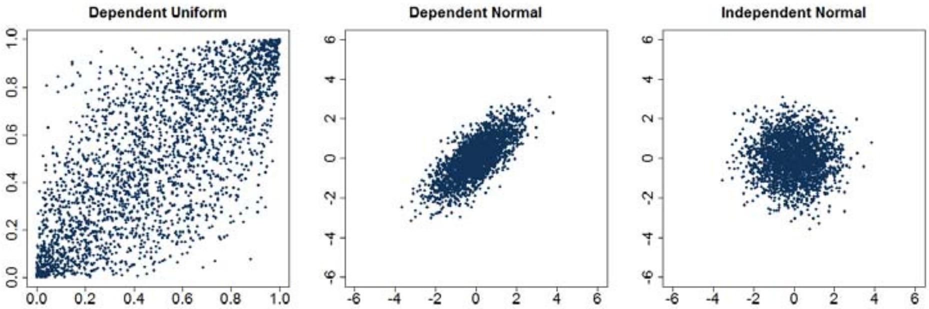
- In practice, Individuated T gives greater ability to target prior beliefs
- But the extra complexity isn't warranted statistically

Review fitted copula
An idea – Group T copulas

Copula	Rho	Nu1	Nu2	Nu3	log L	AIC	D (>3.84)
T	$\begin{bmatrix} 1 & & \\ -0.484 & 1 & \\ 0.159 & -0.281 & 1 \end{bmatrix}$	4.440	-	-	44.21	-80.42	-
Group T (IT with Nu1=Nu2)	$\begin{bmatrix} 1 & & \\ -0.463 & 1 & \\ 0.151 & -0.287 & 1 \end{bmatrix}$	**2.733**	**2.733**	11.391	45.79	-81.57	3.15
Group T (IT with Nu1=Nu3)	$\begin{bmatrix} 1 & & \\ -0.479 & 1 & \\ 0.154 & -0.278 & 1 \end{bmatrix}$	**5.055**	3.191	**5.055**	44.52	-79.05	0.63
Group T (IT with Nu2=Nu3)	$\begin{bmatrix} 1 & & \\ -0.482 & 1 & \\ 0.158 & -0.284 & 1 \end{bmatrix}$	3.587	**4.861**	**4.861**	44.31	-78.62	0.20

- A compromise could be Group T. Here it minimises AIC, but the change in deviance isn't significant, albeit close.
- From here we focus on Gaussian and Student T

Review fitted copula
Goodness of fit – Gaussian



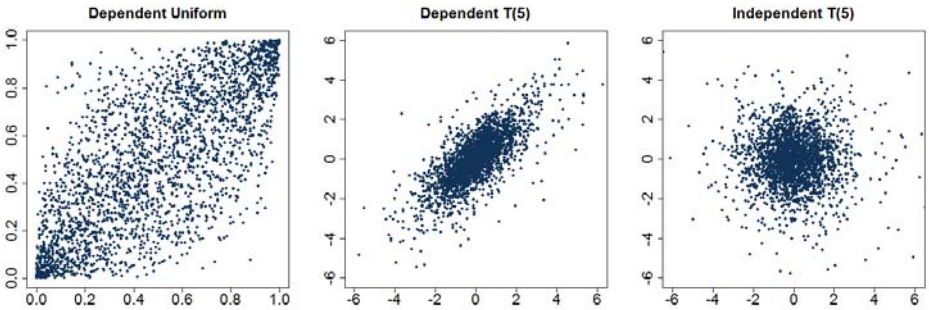
Pseudo-observations (POBS)
$\boldsymbol{U}_i = \begin{pmatrix} u_i \\ v_i \end{pmatrix}$

Normalised POBS
$\boldsymbol{X}_i = \begin{pmatrix} \Phi^{-1}(u_i) \\ \Phi^{-1}(v_i) \end{pmatrix}$
Mardia's test for multivariate normality

Independent
$\frac{1}{\sqrt{1-\rho^2}} \begin{pmatrix} \Phi^{-1}(u_i) - \rho\Phi^{-1}(v_i) \\ \Phi^{-1}(v_i) - \rho\Phi^{-1}(u_i) \end{pmatrix}$
$(\boldsymbol{X}_i - \bar{\boldsymbol{X}})^T \mathrm{S}^{-1} (\boldsymbol{X}_i - \bar{\boldsymbol{X}})$
$\sim$ Scaled β and $\sim \chi^2_d$ as $n \to \infty$

29

Review fitted copula
Goodness of fit – T



Pseudo-observations (POBS)
$\boldsymbol{U}_i = \begin{pmatrix} u_i \\ v_i \end{pmatrix}$

Studentised POBS
$\boldsymbol{X}_i = \begin{pmatrix} t_\nu^{-1}(u_i) \\ t_\nu^{-1}(v_i) \end{pmatrix}$

Independent
$\frac{1}{d}(\boldsymbol{X}_i - \bar{\boldsymbol{X}})^T \mathrm{S}^{-1} (\boldsymbol{X}_i - \bar{\boldsymbol{X}}) \sim F_{d,\nu}$
as $n \to \infty$

30

Review fitted copula

Testing χ^2 and F

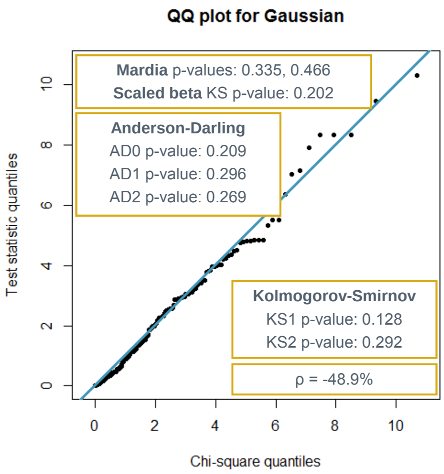
- Under H_0: Gaussian $\mathrm{TS} = (\boldsymbol{X}_i - \overline{\boldsymbol{X}})^T \mathrm{S}^{-1} (\boldsymbol{X}_i - \overline{\boldsymbol{X}}) \sim \chi_d^2$ as $n \to \infty$
- Under H_0: $t(\nu)$ $\mathrm{TS} = \frac{1}{d} (\boldsymbol{X}_i - \overline{\boldsymbol{X}})^T \mathrm{S}^{-1} (\boldsymbol{X}_i - \overline{\boldsymbol{X}}) \sim F_{d,\nu}$ as $n \to \infty$
- We can test these using standard techniques

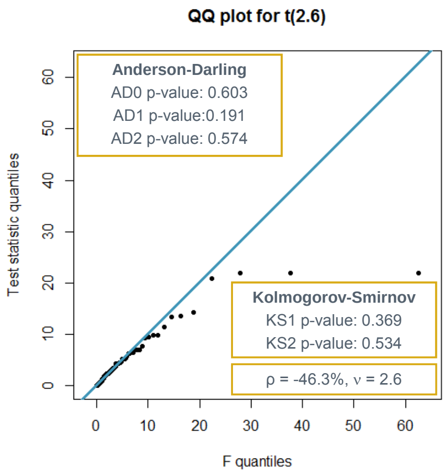
Approach	Anderson-Darling	Kolmogorov-Smirnov
0	$\sum_{i=1}^{n} \frac{1-2i}{n} \left(\log(F_{TS}(x_i)) + \log(1 - F_{TS}(x_i)) \right) - n$	
1	$\max_x \frac{\lvert F_{TS}(x) - F_{TD}(x) \rvert}{\sqrt{F_{TD}(x)(1 - F_{TD}(x))}}$	$\max_x \lvert F_{TS}(x) - F_{TD}(x) \rvert$
2	$\int \frac{\lvert F_{TS}(x) - F_{TD}(x) \rvert}{\sqrt{F_{TD}(x)(1 - F_{TD}(x))}} dF_{TD}(x)$	$\int \lvert F_{TS}(x) - F_{TD}(x) \rvert dF_{TD}(x)$

- Parameters estimated from data $\Rightarrow$ bootstrapping needed

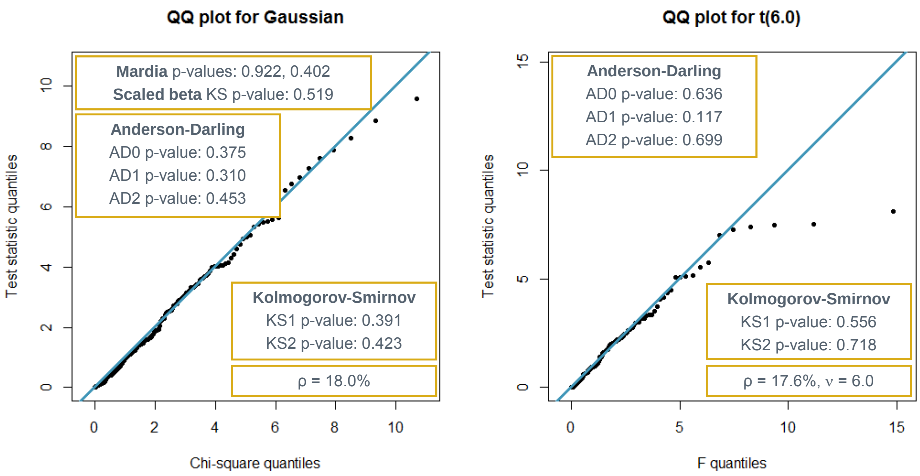
Equity / Credit

Goodness of fit

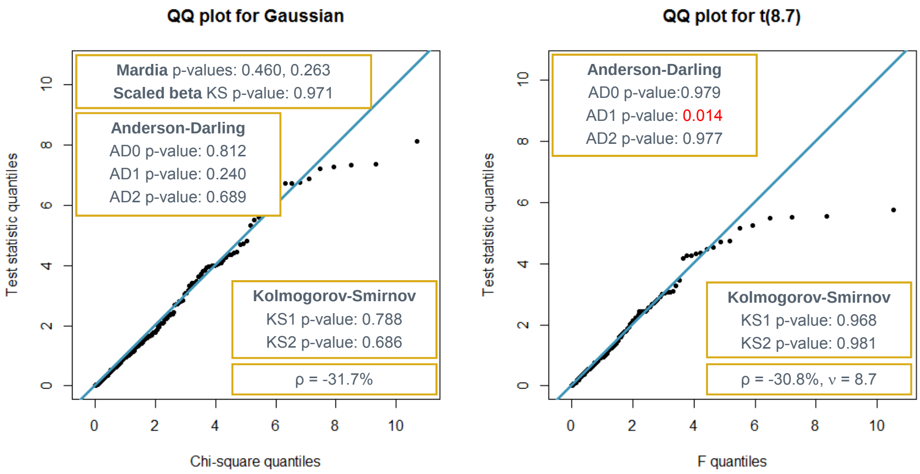




Equity / PC1
Goodness of fit



Credit / PC1
Goodness of fit



Equity / Credit / PC1
Goodness of fit

Copula	Rho	Nu	Anderson-Darling	Kolmogorov-Smirnov
Gaussian	$\begin{bmatrix} 1 & & \\ -0.489 & 1 & \\ 0.183 & -0.318 & 1 \end{bmatrix}$	-	AD0 p-value: 0.146 AD1 p-value: 0.264 AD2 p-value: 0.272	KS1 p-value: 0.161 KS2 p-value: 0.254
T	$\begin{bmatrix} 1 & & \\ -0.484 & 1 & \\ 0.159 & -0.281 & 1 \end{bmatrix}$	4.4	AD0 p-value: 0.685 AD1 p-value: 0.01 AD2 p-value: 0.645	KS1 p-value: 0.806 KS2 p-value: 0.693

Tentative conclusions
Pulling together the various strands

Risk	AIC prefers	D prefers	GoF Gaussian	GoF T
Eq / Cr	T	T	10/10	7/7
Eq / PC1	T	Gaussian	10/10	7/7
Cr / PC1	Gaussian	Gaussian	10/10	6/7
Eq / Cr / PC1	Group T	T	7/7	6/7

- Different measures give different messages
- Gaussian copula fit never rejected on tests applied here
- T copula fit is rarely rejected
- Calls for judgement!

Institute and Faculty of Actuaries

5. Towards a conclusion – the role of judgement

rtise Sponsorship Thought leadership Progress Community Sessional Meetings Education Working parties Volunteering Research Shaping the future Networking Professional support Enterprise and risk Learned society Opportunity International profile Journals Suppo

Choosing and parameterising a copula

Data

- Sparseness or even absence of relevant data
- Range of reasonable values
- Variation over time
- Past not necessarily a guide to the future
- Consistency with risk factor calibration
- Non-coincident time periods can lead to internal inconsistency (non-PSD)

Use Test

- Complexity vs. transparency
- Communication to stakeholders
- Interpretation of parameters
- Sensitivities and alternative assumptions

Expert judgement

- Choice of copula model
- Choice of data and time period
- General reasoning
- Margins to compensate for limitations

Limitations

- Uncertainty
- Symmetry
- Tail dependence

Practical questions

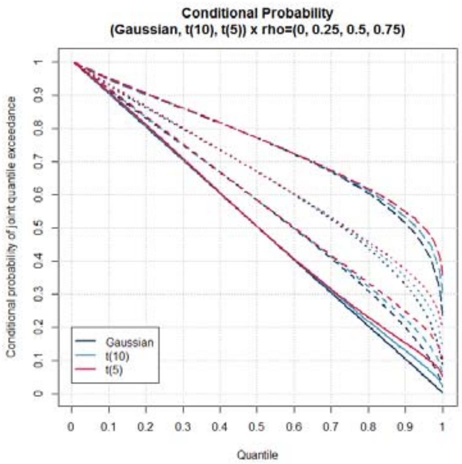
- Can I model it?
- Additional complexity vs. materiality

38

Selecting a copula and assumptions
Tools to aid judgement

- Sensitivity testing (choice of copula, blocks of assumptions, individual assumptions)
  - Standard techniques – not covered further here
- Confidence intervals (parametric or non-parametric)
- Conditional probabilities

Selecting a copula and assumptions
Allowing for tail dependence (or lack of it)



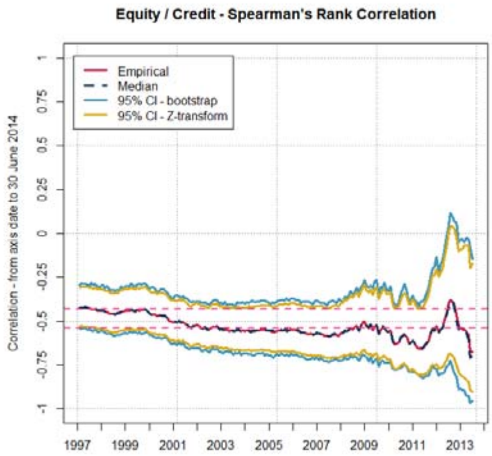
- Practical choice comes down to Gaussian, Student T and Individuated T
- Correlation is primary factor in determining conditional probabilities
- Tail dependence less significant where biting scenario is closer to body of distribution
- If using Gaussian, can make implicit allowance by margins in correlations and validate by sensitivity testing
  - e.g. compare Gaussian with correlation margins to Student T with no margins
- Vast majority of UK internal model firms plan to use Gaussian initially

Selecting a copula and assumptions
Confidence intervals for correlations

- Consider uncertainty surrounding correlation point estimate.
- Bootstrap confidence interval most robust
- Fisher Z CI good approx.

$$\tanh\left(\operatorname{arctanh}\rho \pm \frac{1.96}{\sqrt{n-3}}\right)$$

- Choose in [-54%,-43%]?
  - passes through all CIs
- Additional margins?
- Care needed: PSD

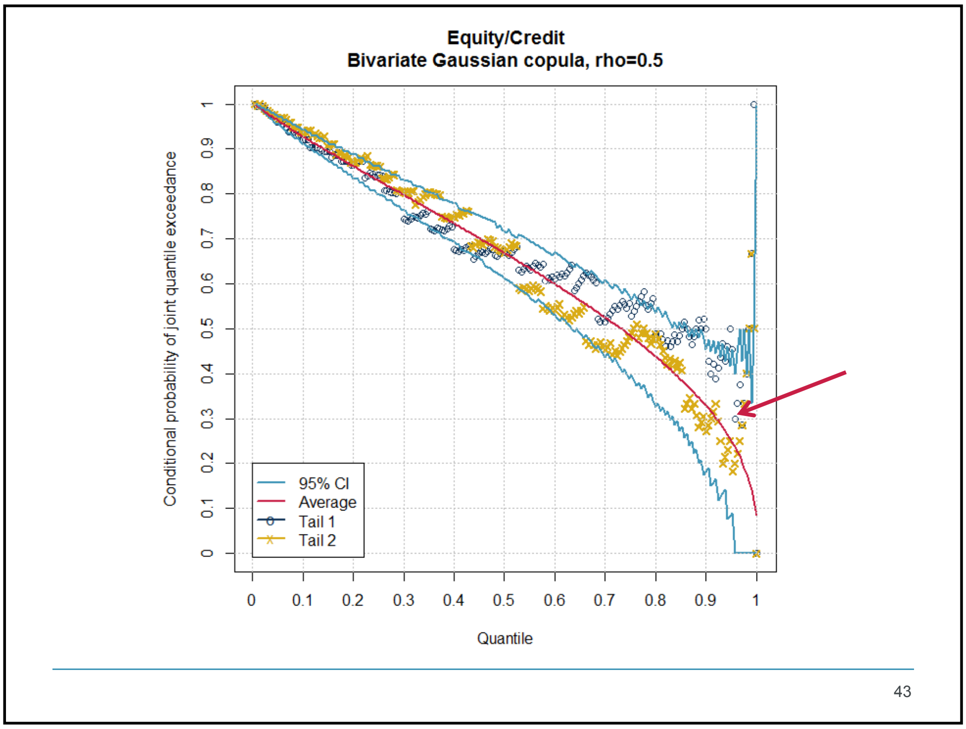


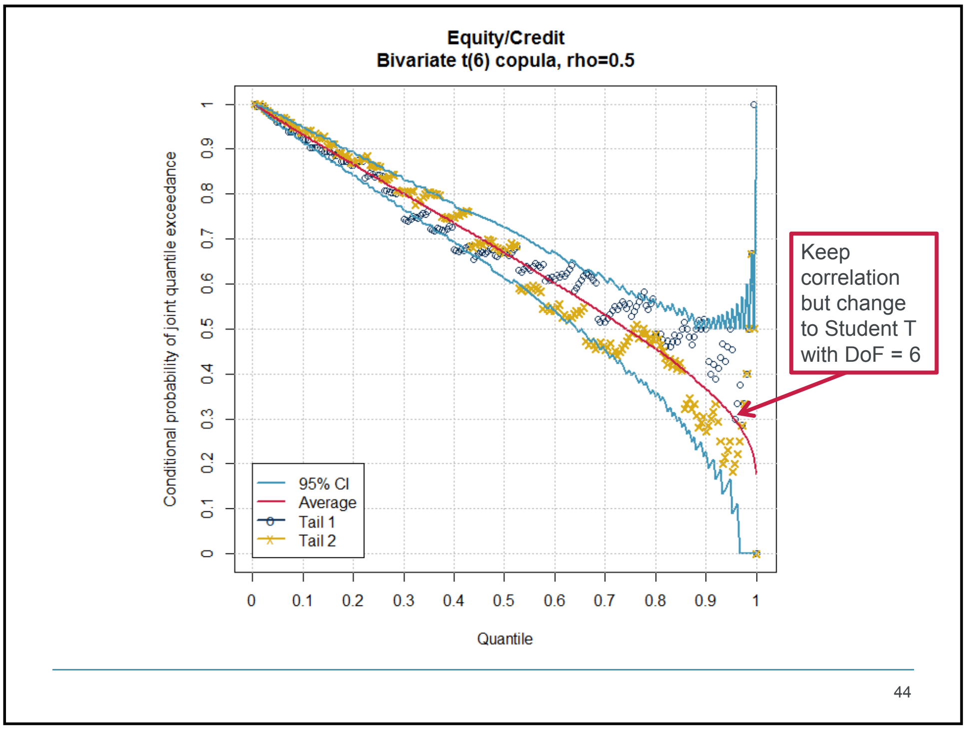
41

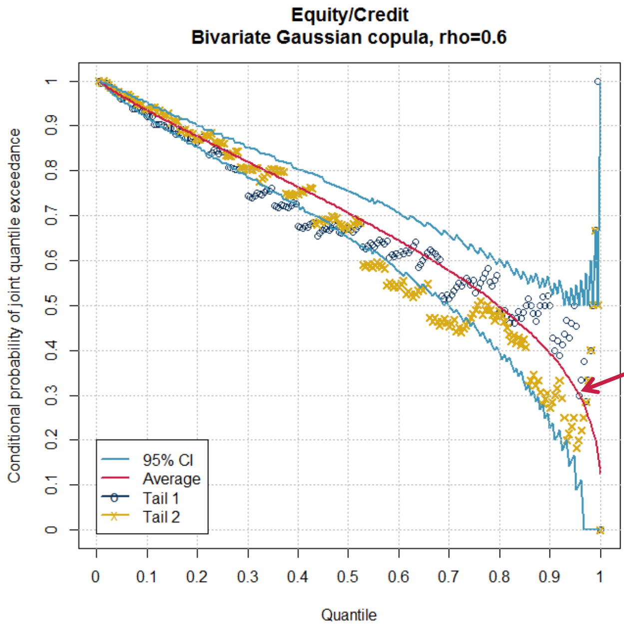
Selecting a copula and assumptions
Conditional probabilities

- Compare empirical conditional probabilities in upper and lower tails with those from chosen copula
- Can then adjust parameterisation of copula to target specific probability or general shape of tail
  - guide choice of correlation assumption for a Gaussian copula
  - guide choice of correlation assumption and degree of freedom for Student T
- See Gary G. Venter "Quantifying Correlated Reinsurance Exposures with Copulas" (Casualty Actuarial Society, 2004)

42


Equity/Credit
Bivariate Gaussian copula, rho=0.5
Conditional probability of joint quantile exceedance
Quantile
95% CI
Average
Tail 1
Tail 2
43


Equity/Credit
Bivariate t(6) copula, rho=0.5
Conditional probability of joint quantile exceedance
Quantile
95% CI
Average
Tail 1
Tail 2
Keep correlation but change to Student T with DoF = 6
44



Keep Gaussian but change correlation

45

Institute and Faculty of Actuaries

6. Next steps for our work

Next steps for our work

- IT copula not developed here – not warranted statistically. However:
  - this may not always be so, e.g. change in data, or other risk factors, and
  - IT is more flexible, albeit harder to parameterise – discarding may be premature.
- We fitted the IT copula successfully, but there were no easy (or quick!) techniques to test the fit. There are bootstrapping techniques in the literature (e.g. Berg (2007) carries out power tests), but:
  - each bootstrap iteration requires the copula to be refitted to simulated data, and
  - at c.½ hour per IT fit, a single test could take c.20 days!
- Augmenting ESG files is an interesting line of approach.
- Vine copulas also.
- Ultimately, expand our work into a SIAS (or perhaps Sessional Meeting) paper.

47

References

- Daniel Berg: “Copula goodness-of-fit testing: an overview and power comparison” (University of Oslo, 2007)
- E. Kole, K Koedijk and M. Verbeek: “Selecting Copulas for Risk Management” (Journal of Banking and Finance 31(8))
- X. Luo and P. Shevchenko: “The t copula with Multiple Parameters of Degrees of Freedom: Bivariate Characteristics and Application in Risk Management” (Quantitative Finance, November 2009)
- Y. Malvergne and D. Sornette: “Testing the Gaussian Copula Hypothesis for Financial Assets Dependencies” (Quantitative Finance, Vol 3, 2003)
- Stephen J. Mildenhall: “Correlation and Aggregate Loss Distributions With An Emphasis On The Iman-Conover Method” (Casualty Actuarial Society, 2005)
- Gary G. Venter: “Quantifying Correlated Reinsurance Exposures with Copulas” (Casualty Actuarial Society, 2004)

48

Questions

Comments

Expressions of individual views by members of the Institute and Faculty of Actuaries and its staff are encouraged.

The views expressed in this presentation are those of the presenters and not of their employers.

49

Institute
and Faculty
of Actuaries

Appendix: copulas – technical background material

rtise nsorship Thought leadership Progress Community Sessional Meetings Education Working parties Volunteering Research Shaping the future Networking Professional support Enterprise and risk Learned society Opportunity International profile Journals Suppo

Copula – technical definition

- A d-dimensional copula is a multivariate distribution function on $[0,1]^d$ with uniform marginals
- i.e. a function $C: [0,1]^d \to [0,1]$ which satisfies the following conditions:
  - *i.* $C(u_1, \dots, u_d) = 0$ if $u_i = 0$ for any i
  - *ii.* $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$ for all $i \in \{1, \dots, d\}$; $u_i \in [0,1]$
  - *iii.* For all $(a_1, \dots, a_d), (b_1, \dots, b_d) \in [0,1]^d$ such that $a_i \le b_i$
    $$\sum_{i_1=1}^{2} \dots \sum_{i_d=1}^{2} (-1)^{(\sum_{j=1}^{d} i_j)} C(u_{1,i_1}, \dots, u_{d,i_d}) \ge 0$$
    where $u_{i1} = a_i$ and $u_{i2} = b_i$ for all $i \in \{1, \dots, d\}$
- Property (iii) ensures that $\Pr\{(u_1, \dots, u_d) \in [a_1, b_1] \times \dots \times [a_d, b_d]\} \ge 0$
- Example: two-dimensional independence copula – $C(u, v) = uv$

Sklar's Theorem

- Isolates the dependency structure (i.e. copula) and marginal distributions of a multivariate distribution
- Every multivariate distribution can be expressed in terms of a copula and its marginal distributions
  - F a joint distribution function with continuous marginals $F_1, \dots, F_d$, then there is a unique copula C such that $F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$ (##) for each $(x_1, \dots, x_d) \in \mathbb{R}^d$
  - Conversely, given a copula C and continuous univariate distribution functions $F_1, \dots, F_d$, the multivariate distribution F defined by (##) has marginals $F_1, \dots, F_d$
- For a continuous multivariate distribution F, its copula is defined by the mapping
  $$(u_1, \dots, u_d) \in [0,1]^d \overset{C}{\to} F\left(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\right) \in [0,1]$$

Why are copulas useful?

- Calculation of economic capital requirements, e.g. Solvency II SCR using an internal model.
- $SCR = q_{0.995}[L(X_1, \dots, X_d)]$ where q is the quantile function and L is the loss function representing losses over a one year time horizon arising from a change $(X_1, \dots, X_d)$ in risk factors.
- Copulas allow you to simulate the $(X_1, \dots, X_d)$ and produce a full probability distribution forecast of profits and losses (required by internal model standards of Solvency II).
- Copulas provide means of separating the loss function L from the risk factor distribution.
- Risk factor distribution $(X_1, \dots, X_d)$ and loss function L can be updated separately. Not possible with variance/covariance matrix approach which would require stress tests to be re-done using actuarial models.
- L can be estimated using a proxy model (e.g. "curve fitting").
- You may have a view on the marginal distributions $F_1, \dots, F_d$ of the changes in risk factors $X_1, \dots, X_d$ $-$ Sklar's Theorem tells us that all you need to define the joint distribution of $(X_1, \dots, X_d)$ is a copula to "glue" the marginals together.

Why are copulas useful?

- Provided we have an algorithm for simulating values $(u_1, \dots, u_d)$ from the copula, it gives us a recipe for generating simulations of the $(X_1, \dots, X_d)$:

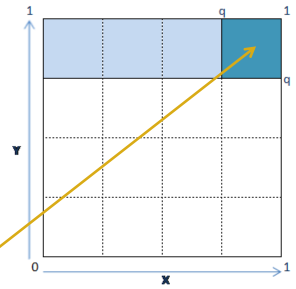
$$(u_1, \dots, u_d) \in [0,1]^d \to (X_1, \dots, X_d) = \left(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)\right)$$

- For example, given a d-dimensional multivariate Normal distribution with mean 0 and correlation matrix Σ with Cholesky decomposition $\Sigma = AA^T$ with A lower triangular:
    i. Generate a d-tuple of independent uniform RVs $U^T = (u_1, \dots, u_d)$ using a standard pseudo-random-number-generator
    ii. Define $Z_i = \Phi^{-1}(U_i)$
    iii. Set $Y = AZ$, Y is a multivariate normal distribution with mean 0 and correlation matrix $AA^T = \Sigma$
    iv. Define $u_j = \Phi(y_j)$
- To assign probabilities and allow for the effects of diversification, need to generate the joint distribution of changes in risk factors $(X_1, \dots, X_d)$. Diversification depends on the risk exposures of the company, granularity of presentation, as well as choice of dependency structure and its parameterisation – can range from 40% to 60%.

Tail dependence – definition

- **Coefficients of upper and lower tail dependence**
- Limiting value of **conditional probability** that extreme value in one variable occurs **given** that an extreme value of the other variable has been realised
- $\lambda_U = \lim_{q \to 1-} \Pr(F_X(X) > q | F_Y(Y) > q) = \lim_{q \to 1-} \frac{C(1-q,1-q)}{1-q}$
- $\lambda_L = \lim_{q \to 0+} \Pr(F_X(X) < q | F_Y(Y) < q) = \lim_{q \to 0+} \frac{C(q,q)}{q}$

Limiting value of ratio of darker shaded to lighter shaded box



Copula	Coefficient of tail dependence
Gaussian	$\lambda_L = \lambda_U = 0$
Student T	$\lambda_L = \lambda_U = 2t_{\nu+1}\left[-\sqrt{(\nu+1)(1-\rho)/(1+\rho)}\right]$
Individuated T	$\lambda_L = \lambda_U = \Omega(\rho, v_1, v_2) + \Omega(\rho, v_2, v_1)$ *Formula for Ω complicated! – see Luo & Shevchenko*
Clayton	$\lambda_L = 2^{-1/\theta}$ for $\theta > 0$, $\lambda_U = 0$
Gumbel	$\lambda_L = 0$, $\lambda_U = 2 - 2^{1/\theta}$ for $\theta > 1$

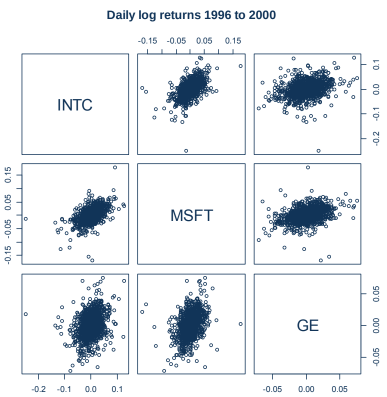
55

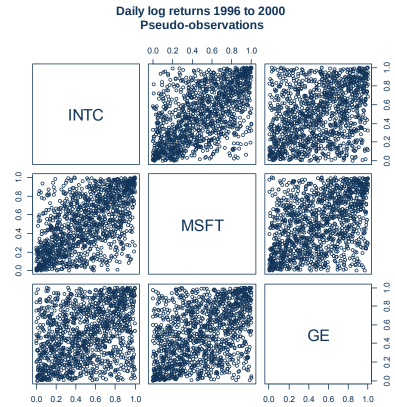
Institute and Faculty of Actuaries

Appendix: textbook example

Theory vs. practice
Textbook example

- Log returns on Intel, Microsoft & General Electric shares over 1996 to 2000 (5 years)
- Daily data





Theory vs. practice
Model postulation

- Raw data suggests an elliptic copula (e.g. Gaussian or Student T)
- Pseudo-observations suggest clustering in tails

Copula	Rho	Nu	log L	AIC
Gaussian	$\begin{bmatrix} 1 & & \\ 0.578 & 1 & \\ 0.340 & 0.402 & 1 \end{bmatrix}$	-	375.5	-745.4
Student T	$\begin{bmatrix} 1 & & \\ 0.588 & 1 & \\ 0.359 & 0.422 & 1 \end{bmatrix}$	6.50	419.3	-830.5

- Likelihood ratio test has p-value of 0 $\Rightarrow$ additional degree of freedom parameter of Student T distribution is significant
- Akaike Information Criterion (AIC) favours Student T

Theory vs. practice

Goodness of fit

Copula	P-value of KS statistic	P-value of Cramer von Mises statistic	Accept H0?
Gaussian	0.00	0.03	Reject
T (DoF=6)	0.26	0.38	Do not reject
T (DoF=7)	0.15	0.45	Do not reject

- Gaussian copula fails goodness of fit test
- Student T copula appears acceptable

Institute
and Faculty
of Actuaries

Appendix: maximum likelihood

Parameter estimation
Maximum likelihood – a reminder

- Choose parameters ϑ to maximise the likelihood $\mathrm{L}(\vartheta) = \prod_{i=1}^{n} g(x_i; \vartheta)$
  - g is the density function
  - $x_i = (x_{i,1}, \ldots, x_{i,d})$ is the observed data
  - $\vartheta = (\vartheta_1, \ldots, \vartheta_p)$ are the parameters to be estimated
  - assuming that the x_i are iid sample
- Maximising the log-likelihood $\ell(\vartheta) = \sum_{i=1}^{n} \log(g(x_i; \vartheta))$ is equivalent

61

Parameter estimation
Maximum likelihood

- We use the result that:

$$C(u_1, \ldots, u_d) = F\left(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\right)$$

$$\Rightarrow\ c(u_1, \ldots, u_d) = \frac{\partial^d C(u_1, \ldots, u_d)}{\partial u_1 \ldots \partial u_d} = \frac{f(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d))}{f_1(F_1^{-1}(u_1)) \ldots f_d(F_d^{-1}(u_d))}$$

to derive the log-likelihood as:

$$\sum_{i=1}^{n} \log\left[f\left(F_1^{-1}(u_{i,1}), \ldots, F_d^{-1}(u_{i,d})\right)\right] - \sum_{i=1}^{n} \sum_{j=1}^{d} \log\left[f_j\left(F_j^{-1}(u_{i,j})\right)\right]$$

- Sometimes some terms are irrelevant to the optimisation
  - e.g. the second term is irrelevant for Gaussian

62

Institute
and Faculty
of Actuaries

63