Bayesian Model Averaging for Estimation of Tail Dependence in Extreme Loss Distributions

Dr Adrian O'Hagan

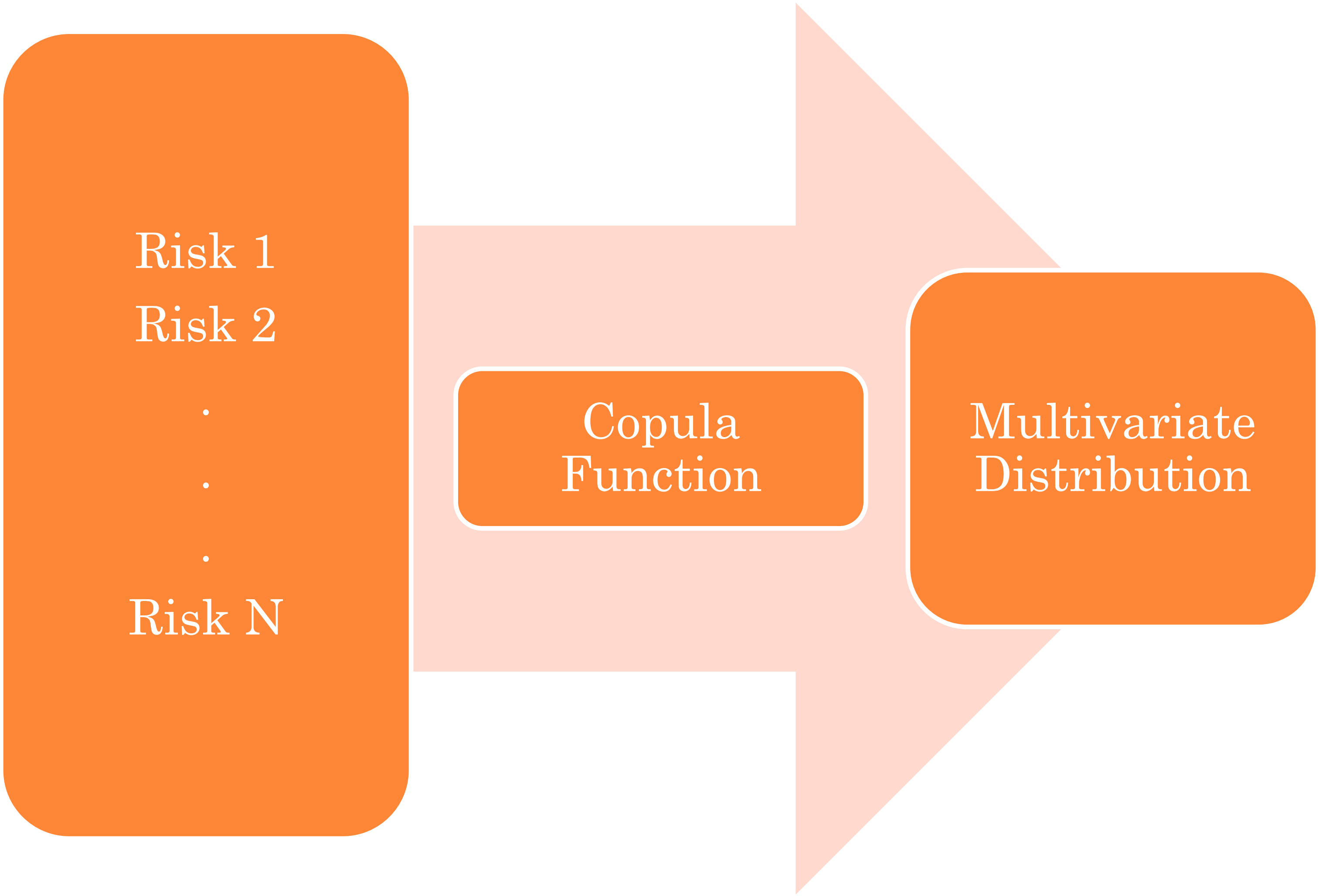
Actuarial Teachers and Researchers' Conference

Edinburgh, December 2014

INTRODUCTION

- OBJECTIVE
  - Assess tail dependency between "random variables"
- APPROACH
  - Copulas
  - Upper tail dependence coefficient
  - Bayesian model averaging
- RESULTS
  - Simulated loss data

COPULA FUNCTIONS



SKLAR'S THEOREM

- For any joint distribution function, $F(x_1, \dots\dots, x_n)$ where each individual risk has its own marginal distribution, $F_{X_i}(x_i)$ …..
- ….. there exists a function C, such that

$$\boldsymbol{F(x_1, \dots\dots, x_n) = C(F_{X_1}(x_1), \dots\dots, F_{X_n}(x_n))}$$

Provided the individual marginal distributions are continuous, the function C is unique.

SELECTED COPULAS

- The t copula
  - Natural successor to the Gaussian copula (?)
  - Incorporates symmetric upper and lower tail dependence.

- The Gumbel and Joe Copulas
  - Incorporate upper tail dependence.
  - Both have lower tail dependence coefficient of 0.

All available through the **copula** package in R.

UPPER TAIL DEPENDENCE COEFFICIENT

$$\lambda_U(X,Y) = \lim_{u \to 1} P(Y > F_Y{}^{-1}(u) \mid X > F_X{}^{-1}(u))$$

and it can be shown that

- $\lambda_U(X,Y) = \lim_{u \to 1}\left(\frac{1 - 2u + C(u,u)}{1 - u}\right)$

using the copula function C

APPROACH

- 1) Simulate loss data.
- 2) Fit chosen copulas to the data.
- 3) Calculate the upper tail dependence coefficient estimate for the data from each copula.
- 4) Weight across the upper tail dependence coefficient estimates.

WEIGHTING ACROSS COPULAS

- Likelihood-based weights
  - Likelihood an obvious gauge of the quality of fit for each copula, and in turn the weight that should be given to its estimated upper tail dependence coefficient.
- $$W_{C_j} = \frac{L_j}{\sum_{i=1}^{n} L_i}$$
- Drags overall estimate of λ in direction of the dominant (highest log-likelihood) copula….
- … but effect less pronounced than for Bayesian Model Averaging weights

WEIGHTING ACROSS COPULAS

- Bayesian Model Averaging based weights
  - $$W_{C_j} = \frac{e^{-\frac{1}{2}BIC_j}}{\sum_{i=1}^{n} e^{-\frac{1}{2}BIC_i}}$$

where BIC_j is the BIC associated with the j^{th}copula

- BIC is the Bayesian information criterion
  - $$BIC = -2lnL + kln(n)$$

- k is the number of parameters
- n is the number of observations

SIMULATED DATA

- Bivariate t distribution using real claims data
  - Simulate data from a t copula based on real claims data.
  - Hence the t copula should provide the best fit.
  - The t copula should give most accurate estimate of λ.

- Bivariate simulated gamma and beta distribution
  - Simulate data from a t copula based on underlying simulated claims data.
  - The t copula should again provide the best fit.
  - The t copula should again give most accurate estimate of λ.

TRUE VALUE OF UPPER TAIL DEPENDENCE COEFFICIENT FOR T COPULA

$$\lambda_{true} = 2t_{v+1}\{(-\sqrt{v+1}\sqrt{1-\rho})/\sqrt{1+\rho})\}$$

- v is the degrees of freedom
- ρ is the correlation coefficient
- t_{v+1}(k) represents $P(t_{v+1} < k)$

SIMULATED DATA

- Bivariate t data: $\lambda_{true} = 0.232$

Copula	Upper Tail Dependence Coefficient	BIC
t	0.238	-1,205.35
Gumbel	0.471	-146.74
Joe	0.620	-157.58

- Bivariate gamma and beta data: $\lambda_{true} = 0.785$

Copula	Upper Tail Dependence Coefficient	BIC
t	0.781	-1,439.96
Gumbel	0.764	-1,435.68
Joe	0.759	-1431.34

RESULTS: BIVARIATE T DATA

The BIC for the "best" (correct) model is significantly greater than that for the other models.

- $$e^{-\frac{1}{2}BIC_{C_j}} \approx 0 \quad \forall\, C_j \neq \text{the t copula}$$
- Hence in this case the BMA weighting approach simply identifies the correct model as giving the best estimate of upper tail dependence:
- $$\lambda_{weighted} = \lambda_t \approx \lambda_{true}$$

RESULTS: BIVARIATE GAMMA AND BETA DATA

- The BICs for all candidate models are relatively close.

- Using Likelihood-based weights:
- $W_{C_t} = 0.335$ $W_{C_{Gumbel}} = 0.333$ $W_{C_{Joe}} = 0.332$
- $\lambda_{true} = 0.785$ $\lambda_{Likelihood} = 0.768$

- Using BMA:
- $W_{C_t} = 0.884$ $W_{C_{Gumbel}} = 0.104$ $W_{C_{Joe}} = 0.012$
- $\lambda_{true} = 0.785$ $\lambda_{BMA} = 0.779$

CONCLUSIONS

- Bayesian model-averaging provides a computationally straightforward, statistically robust way to:

- 1) Identify when a copula model for tail dependence is significantly better than other candidates.

OR

- 2) Blend information from multiple copula models for tail dependence when more than one model is “good”.

FURTHER WORK

- R package **BMAcopula**
  (for absorption into the **copula** package)

- Research paper
  (paired with empirical copula tail dependence
  coefficient estimation)

REFERENCES

- "*Measurement and modelling of dependencies in economic capital, a discussion paper*". Shaw, Smith & Spivak, May 2010.
- "*The t copula and related copulas*". Demarta & McNeil, May 2004.
- "*Bayesian model averaging in R*". Amini & Parmeter.
- "*Modelling the dependence structure of financial assets: a survey of four copulas*". Aas, Dec 2004.