Correlation & Dependency Structures

GIRO - October 1999

Andrzej Czernuszewicz
Dimitris Papachristou
Why are we interested in correlation/dependency?

• Risk management
• Portfolio management
• Reinsurance purchase
• Pricing
How does dependency arise in the insurance world?

Related Events

- insurance cycle
- economic factors
- physical events
- social trends
- reinsurance failure
Items to be covered:

1. Simulations of correlated variables
2. Problems with the traditional approach
3. Copulas
Statement of problem

Suppose we have $n$ classes of business with each class of business having its own marginal distribution. How do we model the portfolio?

Traditionally this problem is tackled using correlation as the measure of dependence.
• The mean does not tell everything about a distribution.

• The coefficient of correlation does not tell everything about the dependency structure.
Standard Simulation Technique

Step 1

Simulate an $n$-dimensional multivariate normal distribution with correlation matrix $\rho$
Simulate Multivariate Normal Variables

Simulated Independent Normal

Simulated Multivariate Normal
Standard Simulation Technique

Step 2
generate $n$ series of the appropriate marginal distributions and put into a matrix

Step 3
Apply Normal multivariate dependency structure from step 1
Standard Simulation Technique

**Simulated Independent Normal**

**Simulated Multivariate Normal**

**Simulated Independent Gamma Variables**
Observations

- The Spearman rank correlations are matched rather than the Pearson correlations
- The solution is not unique
- In particular use of normal distribution influences the tail of the modelled portfolio
The same Correlation, but Different Dependency Structures

Number of points in upper quartile = 9

Number of points in upper quartile = 23
Observation

In the majority of applications a symmetric approach is used in determining the dependency. However in practice this need not be the case.

Example: An earthquake may cause a stockmarket crash but a stockmarket crash will not cause an earthquake.

Skewed distributions may be used to get around this problem.
Asymmetric Dependency Structure

The diagram shows the number of points in the upper quartile as 37.
The same Correlation, but Different Dependency Structures
Two Fallacies

Fallacy 1
Marginal distributions and correlations determine the joint distribution

Fallacy 2 (not for rank correlation)
F_1, F_2 marginal distribution for X_1, X_2

\[ \forall \rho \in [-1,1] \  \exists \ F \text{ such that } F \text{ is the joint distribution and } X_1, X_2 \text{ have Pearson correlation } \rho \]
Other problems with Pearson correlation

**Problem 1**
- A correlation of zero does not indicate independence of risk.

**Problem 2** (not for rank)
- Correlation is not invariant under transformations of the risk.

**Problem 3**
- Correlation is not an appropriate dependence measure for very heavily-tailed distributions.
Why dependency structures?

We need to amend our concept of dependency to allow for desirable features

- In particular we have to introduce non-linear dependency. For example:
  \[ X \sim u[-1,1], \quad Y = X^2, \quad \rho = 0 \]

- Need to reflect special features of tail dependence.
- In general, single numeric measures of dependency are insufficient.
The Copula approach

• Operates by separating marginal distributions from dependency structures

• Combination of copula and marginal will yield original distribution exactly

• No longer have the problems associated with correlation
Definition of Copula

For m-variate distribution F with j th univariant margin F_j
the copula associated with F is a distribution function

\[ C : [0,1]^m \rightarrow [0,1] \text{ that satisfies } \]
\[ F(X) = C(F_1(X_1),...,F_m(X_m)) \]

(Note: if F is a continuous m-variate distribution the copula
associated with F is unique)

This can also be represented via a density function

\[ c(u,v) = \frac{\partial C(u,v)}{\partial u \partial v}, \quad 0 < u, v < 1 \]
Construction of a Copula

• Construction of a copula
  – parametric
  – non-parametric

• parametric form needed for higher dimension problem
Example of Parametric Copula (1)

Independent Copula

\[ C(u, v) = uv, \quad 0 \leq u, v \leq 1 \]
\[ c(u, v) = 1, \quad 0 \leq u, v \leq 1 \]
Example of Parametric Copula (2)

**Normal copula**

\[
c(u, v, \rho) = (1 - \rho^2)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} (1 - \rho^2)^{-1} [x^2 + y^2 - 2\rho xy] \right) \cdot \exp\left\{\frac{1}{2} [x^2 + y^2] \right\}
\]

where

\[x = \Phi^{-1}(u), \quad y = \Phi^{-1}(v)\]

\[\Phi \text{ is } N(0,1)\]

and \(\rho\) is the correlation.
Example of Parametric Copula (3)

**Gumbel Copula**

For $1 \leq \delta < \infty$

$$C(u, v; \delta) = \exp \left\{ - \left( (-\log u)^\delta + (-\log v)^\delta \right)^{\frac{1}{\delta}} \right\} \quad 0 \leq u, v \leq 1$$

Note: this is an extreme value copula
Independent (Product) Copula Density
Simulation of a copula

• simulate a value $u_1$ from $U(0,1)$

• simulate a value $u_2$ from $C_2(u_2 | u_1)$

• simulate a value $u_n$ from $C_n(u_n | u_1, ..., u_{n-1})$

where $C_i = C(u_1, ..., u_i, 1, ..., 1)$ for $i=2, ..., n$
Example: Reinsurance Pricing

- Stop Loss on two classes of business
- Assumptions

<table>
<thead>
<tr>
<th></th>
<th>Class A</th>
<th>Class B</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean L/R</td>
<td>91.14%</td>
<td>91.85%</td>
</tr>
<tr>
<td>st. dev. of L/R</td>
<td>10.98%</td>
<td>9.21%</td>
</tr>
<tr>
<td>Premium</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

- Limit: 15%
- Deductible: 107.5%
- Expected L/R: 91.5%
Loss Ratios are assumed to be lognormally distributed.
The same rank correlation ($\rho = 0.25$), but different dependency structures especially at the tail.
Results

Expected amount of loss to the layer is underestimated approximately by 50%, if a multivariate Normal dependency structure is used.

<table>
<thead>
<tr>
<th></th>
<th>Gumbel</th>
<th>Independent</th>
<th>Mult. Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>0.325</td>
<td>0.116</td>
<td>0.220</td>
</tr>
<tr>
<td>standard deviation</td>
<td>2.297</td>
<td>1.169</td>
<td>1.735</td>
</tr>
<tr>
<td>Rate on Line</td>
<td>1.1%</td>
<td>0.4%</td>
<td>0.7%</td>
</tr>
</tbody>
</table>
Conclusions

• Correlation is not a sufficient measure of dependence.

• Opportunities may be missed by remaining in the correlation framework.

• True dependency reflected in the copula approach by separating marginal distribution from dependency structures.