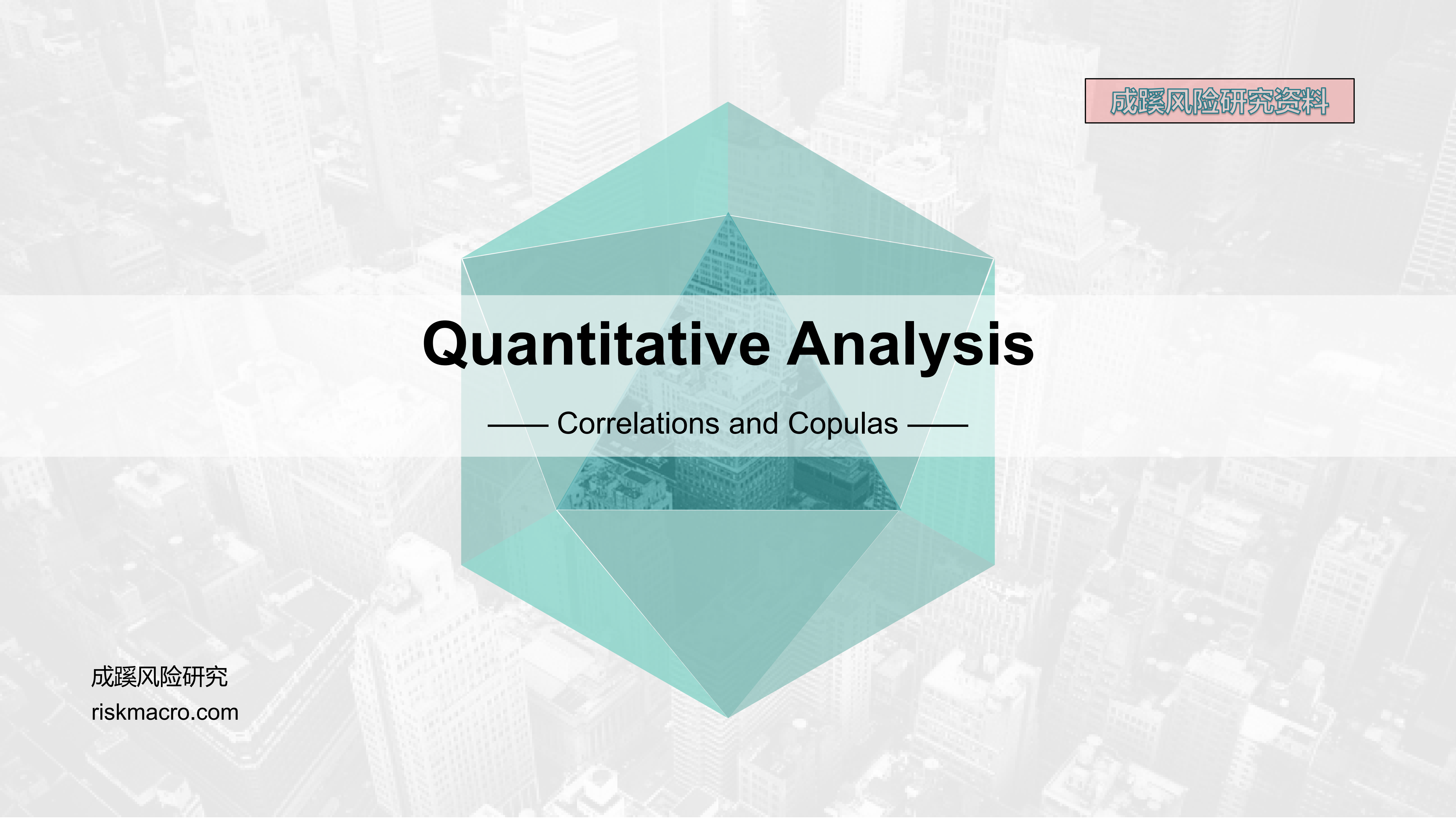


Quantitative Analysis

—— Correlations and Copulas ——

成蹊风险研究
riskmacro.com

Correlations and Copulas

Correlations and Copulas

CORRELATION AND COVARIANCE:

- To calculate the standardized correlation of two variable, we need to
  - 1st calculate the expected value of standard deviations of each variables
  - 2nd calculate the expected value of covariance between variable, then
  - 3rd calculate the correlation between variable

$$E(\sigma_X^2) = E(X^2) - [E(X)]^2$$

$$E(\sigma_Y^2) = E(Y^2) - [E(Y)]^2$$

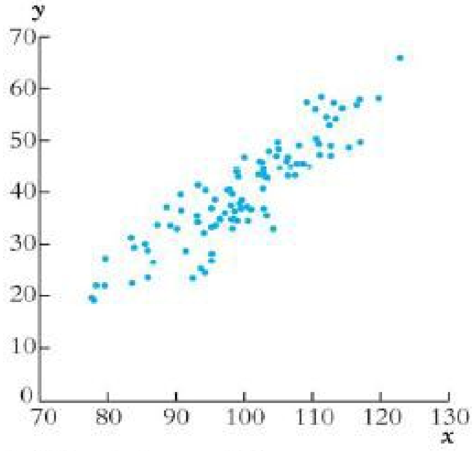
$$E\left(Cov_{X,Y}\right) = E(XY) - [E(X)E(Y)]$$

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_x \sigma_Y}$$

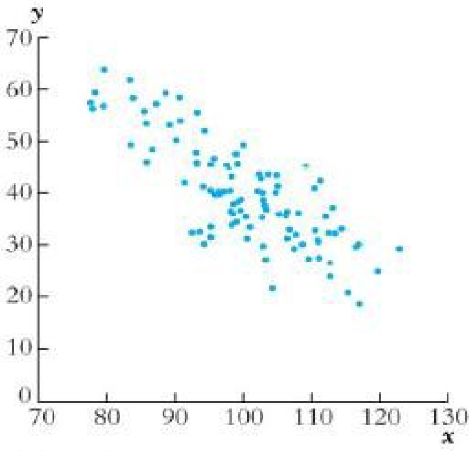
- A correlation of zero implies that, there is no linear relationship between the two variables, but the value of one variable can still have a nonlinear relationship with the other variable.
- The coefficient of correlation lies between -1 to +1.

Correlations and Copulas

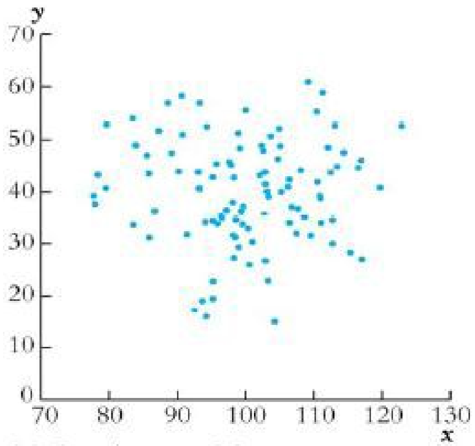
- **CORRELATION AND COVARIANCE:**



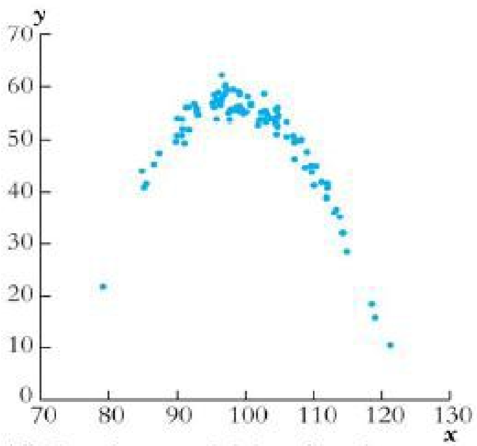
(a) Correlation = +0.9



(b) Correlation = −0.8



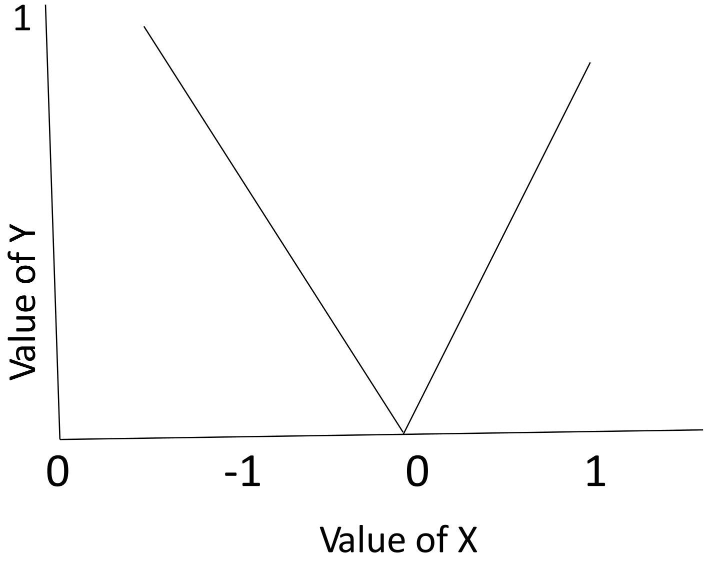
(c) Correlation = 0.0

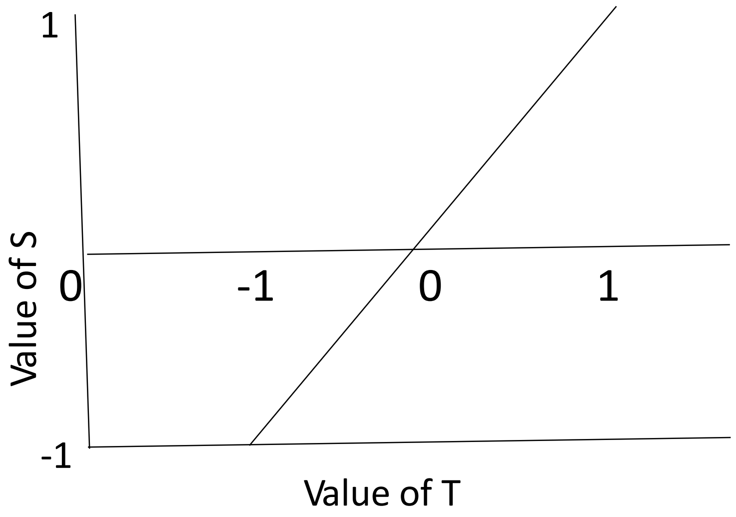


(d) Correlation = 0.0 (quadratic)

Correlations and Copulas

- **Correlation vs. Dependence:**





<u>Example: Non-linear relations</u>

Correlations and Copulas

- **COVARIANCE USING EWMA AND GARCH MODELS:**

☐ EWMA MODEL

The exponentially weighted moving average (EWMA) model is designed to vary the weight given to more recent observations (by adjusting λ).

$$Cov_n = \lambda cov_{n-1} + (1 – \lambda)\ X_{n-1}\ Y_{n-1}$$

Where:

λ =the weight for the most recent covariance on day n-1

X_{n-1} = the percentage change for variable X on day n-1 y_{n-1} = the percentage change for variable Y on day n-1

Correlations and Copulas

- COVARIANCE USING EWMA AND GARCH MODELS:

❑ EWMA MODEL

Example: Calculating covariance using the EWMA model

Assume an analyst uses the EWMA model with λ=0.90 to update correlation and covariance rates. The correlation estimate for two variables X and Y on day n-1 is 0.7. in addition, the estimated standard deviation on day n-1 for variables X and Y are 1.5% and 2% respectively. Also, the percentage change on day n-1 for variables X and Y are 2% and 1%, respectively. What is the updated estimate of the covariance rate and correlation between X and Y on day n?

Correlations and Copulas

COVARIANCE USING EWMA AND GARCH MODELS:

❑ GARCH(1,1) MODEL

Omega

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH)(1,1) model, applies a weight of α to the most recent observation on covariance (X_{n-1} Y_{n-1}) and a weight of β to the most recent covariance estimate (cov_{n-1}). In addition, a weight of ω is given to the long-term average covariance rate.

$$\mathrm{Cov_n} = \omega + \alpha X_{n-1} Y_{n-1} + \beta \mathrm{cov}_{n-1}$$

Where:

α =the weight for the most recent covariance on Covariance (X_{n-1} Y_{n-1})
β = the weight for the most recent covariance estimate (Cov_{n-1})
ω = the weight for the long term average covariance rate

- *The EWMA is a special case of GARCH (1, 1),*

where

w =0, α = 1- λ, and β = λ
ω = ϒ V_L where ϒ is the weight assigned to long term variance V_L , also
α + β + ϒ = 1, hence V_L = ω/ϒ = ω/(1- α – β)

Correlations and Copulas

➢ COVARIANCE USING EWMA AND GARCH MODELS:

❑ GARCH(1,1) MODEL

Example: Calculating covariance using the GARCH(1, 1) model

Assume an analyst uses daily data to estimate a GARCH(1, 1) model as follows:

$$cov_n = 0.000002 + 0.14X_{n-1}Y_{n-1} + 0.76cov_{n-1}$$

this implies α = 0.14, β = 0.76, and and ω = 0.000002. The analyst also determine that the estimate of covariance on day n-1 is 0.018 and the most recent observation on covariance is 0.02. what is the updated estimated of covariance?

Correlations and Copulas

❑ *Vectors*

- Vectors are defined as quantities that have a magnitude (or length) and a direction. They can be thought of as representing the position of a point in two dimensions, three dimensions *etc*.
- All vectors in three-dimensional space can be written in terms of the base vectors **i**, **j** and **k**, which are the unit vectors (*ie* the vectors of length 1) in the *x*, *y* and *z* directions of cartesian space.
- There are different ways of writing vectors:

$$a = \underline{a} = 2i + 3j - 2k = \begin{bmatrix} 2 \\ 3 \\ -2 \end{bmatrix} = \begin{pmatrix} 2 & 3 & -2 \end{pmatrix}^T$$

where $i = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$ $j = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}$ $k = \begin{pmatrix} 0 \\ 0 \\ -3 \end{pmatrix}$

❑ *Magnitude*

- The magnitude of a vector is just its length, which can be calculated using an extended version of Pythagoras' theorem. The magnitude of vector **a** is written as *a* or | **a** | .
- In general, if $a = \begin{pmatrix} p \\ q \\ r \end{pmatrix}$, then $a = \sqrt{p^2 + q^2 + r^2}$

❑ *Matrices*

- Matrices are arrays of numbers whose size is referred to as the number of rows by the number of columns. Notice that a vector is a special case of a matrix where the number of columns is one. A square matrix is one where the number of columns equals the number of rows.
- The transpose of a matrix is found by swapping the rows and the columns. The transpose of a matrix **A** is written as $\mathbf{A}^T$ or **A**′ .
- Example:
- If $A = \begin{pmatrix} 1 & -3 \\ 2 & 5 \\ 0 & 9 \end{pmatrix}$, then $A^T = \begin{pmatrix} 1 & 2 & 0 \\ -3 & 5 & 9 \end{pmatrix}$

Correlations and Copulas

❑ *Determinants*

A determinant is a scalar quantity associated with a square matrix. The determinant of a 2 x 2 matrix is equal to the product of the numbers on the leading diagonal (top left corner to bottom right corner) minus the product of the numbers on the other diagonal. It is written as det**A** , | **A** | , or Δ when it is clear which matrix is involved.

Example:

What is det**A** if $A = \begin{pmatrix} 2 & -6 \\ 4 & 3 \end{pmatrix}$

det**A** = (2 x 3) - (4 x -6) = 30

❑ *Inverses*

The inverse of a matrix **A** , written as $\mathbf{A}^{-1}$, is such that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\,\mathbf{A} = \mathbf{I}$. It is the matrix equivalent to a reciprocal.

To find the inverse of a 2 x 2 matrix, swap the elements on the leading diagonal, change the sign of the elements on the other diagonal and divide by the determinant, *i.e.*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Correlations and Copulas

EVALUATING CONSISTENCY FOR COVARINCES:

- The diagonal of the variance , co-variance matrix represents the variances rates where ***i=j***.
- The covariance rates are all other elements of the matrix where ***i≠ j***.
- A matrix is known as positive-semi definite if it is internally consistent.
- The condition for an N x N variance- covariance matrix, to be internally consistent for all N x 1 vectors ω, where ω^T is the transpose of vector ω.

Variance, Co-variance matrix of dimension NxN

Transpose of ω

$$\omega^T \Omega \omega \geq 0$$

Any vector of N real number

- The variance , co-variance matrix must be positive –semi definite.
- Small changes to a small PS matrix will likely be still PS, but small changes to large PS matrix , no longer be PS.

Correlations and Copulas

EVALUATING CONSISTENCY FOR COVARINCES:

- Example

$$\begin{array}{cccccc} & \sigma_1^2 & Cov_{1,2} & Cov_{1,3} & \\ & 1 & 0 & 0.8 & \\ Cov_{2,1} & 0 & \overset{\sigma_2^2}{1} & 0.8 & Cov_{2,3} \\ Cov_{3,1} & 0.8 & 0.8 & \overset{\sigma_3^2}{1} & \\ & & Cov_{3,2} & & \end{array}$$

$$\rho_{1,2} = \frac{Cov_{1,2}}{\sigma_1 \sigma_2} = \frac{0}{1x1} = 0$$

$$\rho_{1,3} = \frac{Cov_{1,3}}{\sigma_1 \sigma_3} = \frac{0.8}{1x1} = 0.8$$

$$\rho_{2,3} = \frac{Cov_{2,3}}{\sigma_2 \sigma_3} = \frac{0.8}{1x1} = 0.8$$

- The matrix is positive definite if all its principal minors have strictly positive determinants.

$$\begin{array}{cccc} \Lambda_1 \geq 0 & 1 & 0 & 0.8 \\ \Lambda_2 \geq 0 & 0 & 1 & 0.8 \\ \Lambda_3 \geq 0 & 0.8 & 0.8 & 1 \end{array}$$

$$\Lambda_1 = 1x1 = 1 \geq 0$$

$$\Lambda_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1x1 - 0x0 = 1 \geq 0$$

$$\Lambda_3 = 1x \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix} - 0x \begin{bmatrix} 0 & 0.8 \\ 0.8 & 1 \end{bmatrix} + 0.8x \begin{bmatrix} 0 & 1 \\ 0.8 & 0.8 \end{bmatrix}$$

$$= 1x(1 - 0.64) - 0 + 0.8x(0 - 0.8)$$

$$= 0.36 - 0.64 = -0.28 \ngeq 0$$

- Hence the above matrix is not positive semi-definite.

Correlations and Copulas

EVALUATING CONSISTENCY FOR COVARINCES:

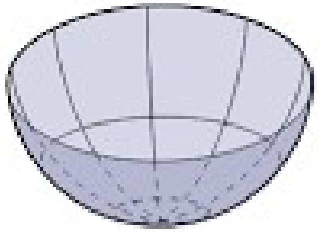
- **Other method for testing for consistency:**

$$\rho_{12}^2 + \rho_{13}^2 + \rho_{23}^2 - 2_{\rho_{12}\rho_{13}\rho_{23}} \leq 1$$

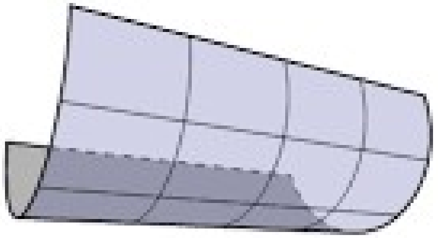
- Example: Check internal consistency of the matrix

$$0^2 + 0.8^2 + 0.8^2 - 2x0x0.8x0.8 = 1.28$$

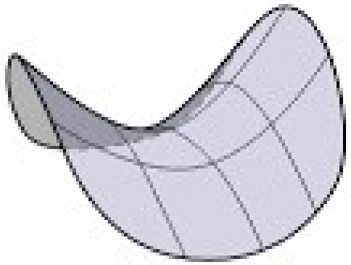
- Indicating that the matrix is not internally consistent.



Definite



Semi-definite



Indefinite

Correlations and Copulas

GENERATING SAMPLES :

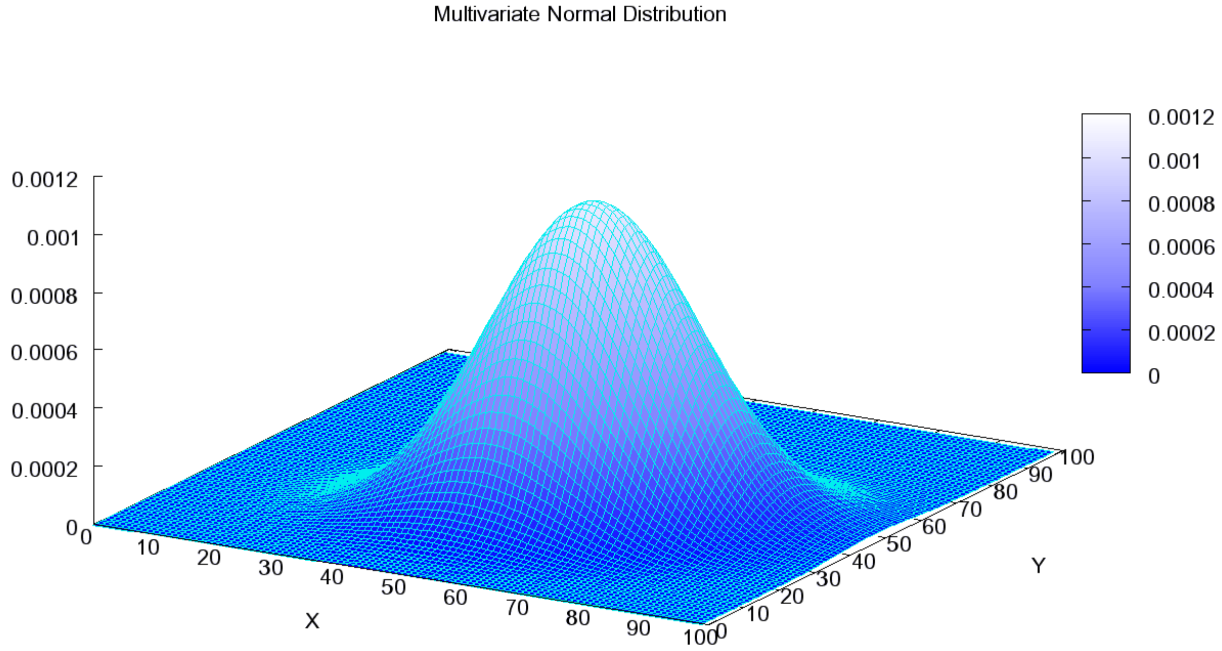
- In a bivariate normal distribution between variable X & Y, the expected mean of dependent variable Y is:
- $\mu_Y + \beta(X - \mu_X)$ where $\beta = \frac{Cov_{XY}}{\sigma_X^2} = \frac{\rho_{XY}\sigma_Y}{\sigma_X}$
- And standard deviation of Y is:

$$\sigma_Y\sqrt{1 - \rho_{XY}^2}$$

- The conditional sample of Y variable is:

$$\varepsilon_Y = \rho_{XY}Z_X + Z_Y\sqrt{1 - \rho_{XY}^2}$$

- Where Z_x and Z_y are independent samples, obtained from a univariate standardized normal distribution.



Correlations and Copulas

- **FACTOR MODELS:**
  - A factor model can used to define correlations between normally distributed variables.
  - The following equation is a one-factor model where each U_i has a component depend on one common factor(F) in addition to another component (Z_i) that is uncorrelated with other variables.

$$U_i = \alpha_i F + Z_i \sqrt{1 - \alpha_i^2}$$

  - Every U_i has a standard normal distribution (mean =0, standard deviation =).
  - The constant α_i is between -1 and 1.
  - F and Z_i have standard normal distribution and are uncorrelated with each other.
  - Every Z_i is uncorrelated with each other.
  - All correlations between U_i and U_j result from their dependence on a common factor, F.

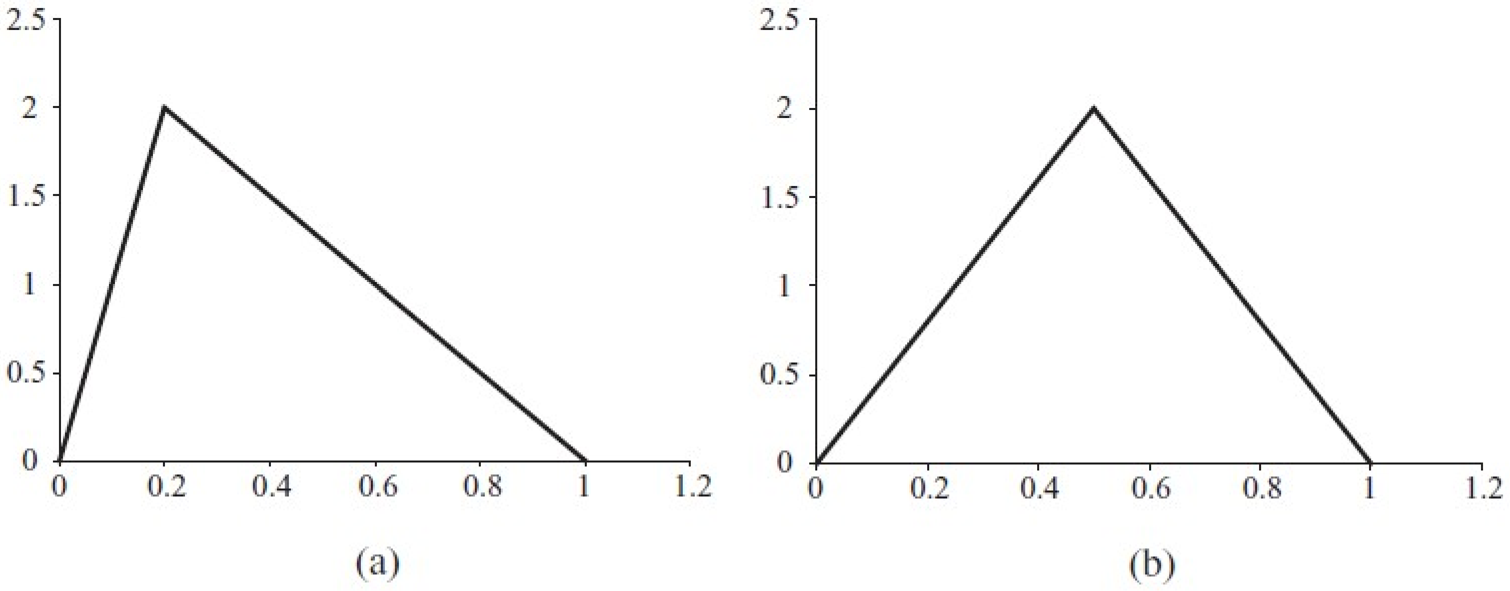
Advantage of one factor Model

  - the covariance matrix for a one-factor model is positive-semi-definite.
  - the number of correlations between variables is greatly reduced.

- CAPM model is one factor model.

Correlations and Copulas

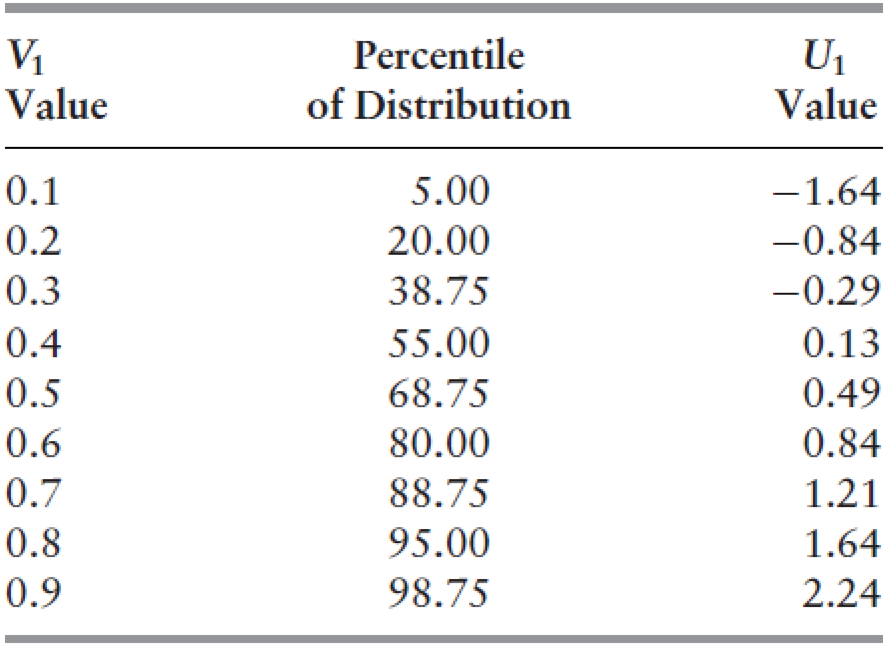
- **COPULAS:**
  - A copula create a joint probability distribution between two or more variable while maintaining their individual marginal distribution.
  - This is accomplished by mapping the marginal distribution to a new known distribution.
  - Gaussian copula maps the marginal distribution of each variable to the standard normal distribution, which, by definition, has a mean of zero and a standard deviations of one.
  - The mapping of each variance to the new distributions done based on percentiles.



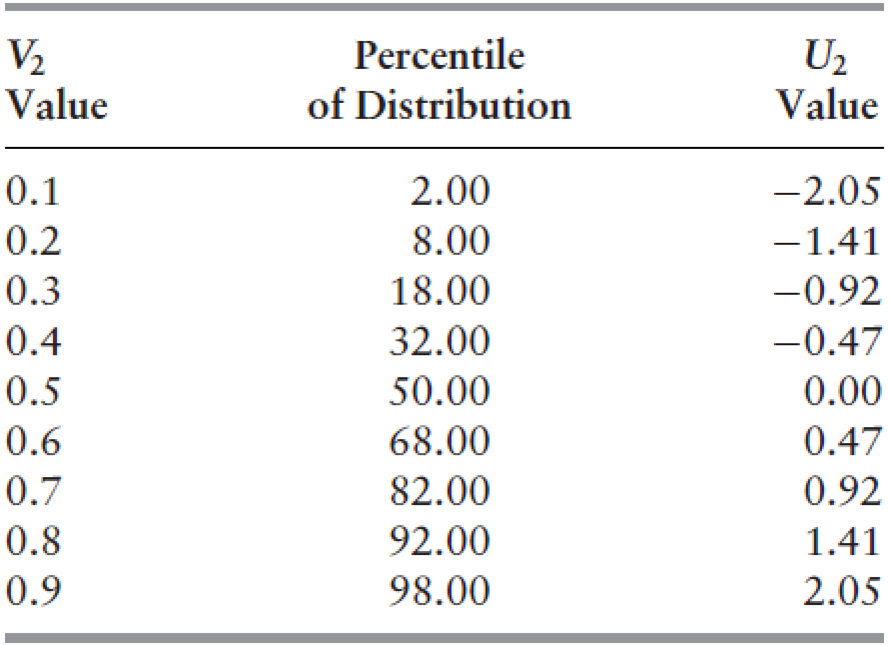
Triangular Distributions for V_1 and V_2

Correlations and Copulas

- **COPULAS:**
  - For example



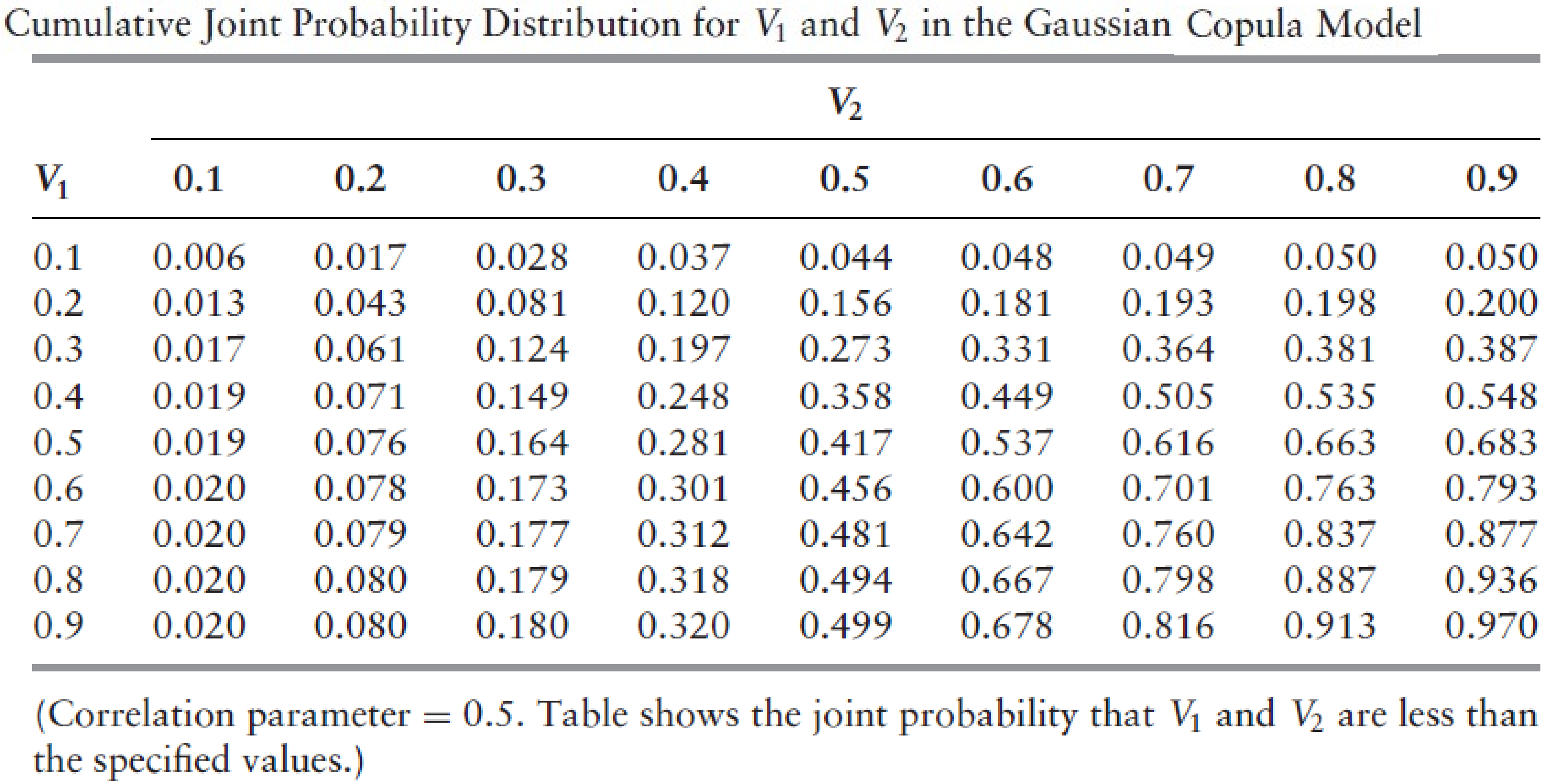
V_1 Value	Percentile of Distribution	U_1 Value
0.1	5.00	−1.64
0.2	20.00	−0.84
0.3	38.75	−0.29
0.4	55.00	0.13
0.5	68.75	0.49
0.6	80.00	0.84
0.7	88.75	1.21
0.8	95.00	1.64
0.9	98.75	2.24



V_2 Value	Percentile of Distribution	U_2 Value
0.1	2.00	−2.05
0.2	8.00	−1.41
0.3	18.00	−0.92
0.4	32.00	−0.47
0.5	50.00	0.00
0.6	68.00	0.47
0.7	82.00	0.92
0.8	92.00	1.41
0.9	98.00	2.05

Correlations and Copulas

- **COPULAS:**
  - For example



Cumulative Joint Probability Distribution for V_1 and V_2 in the Gaussian Copula Model

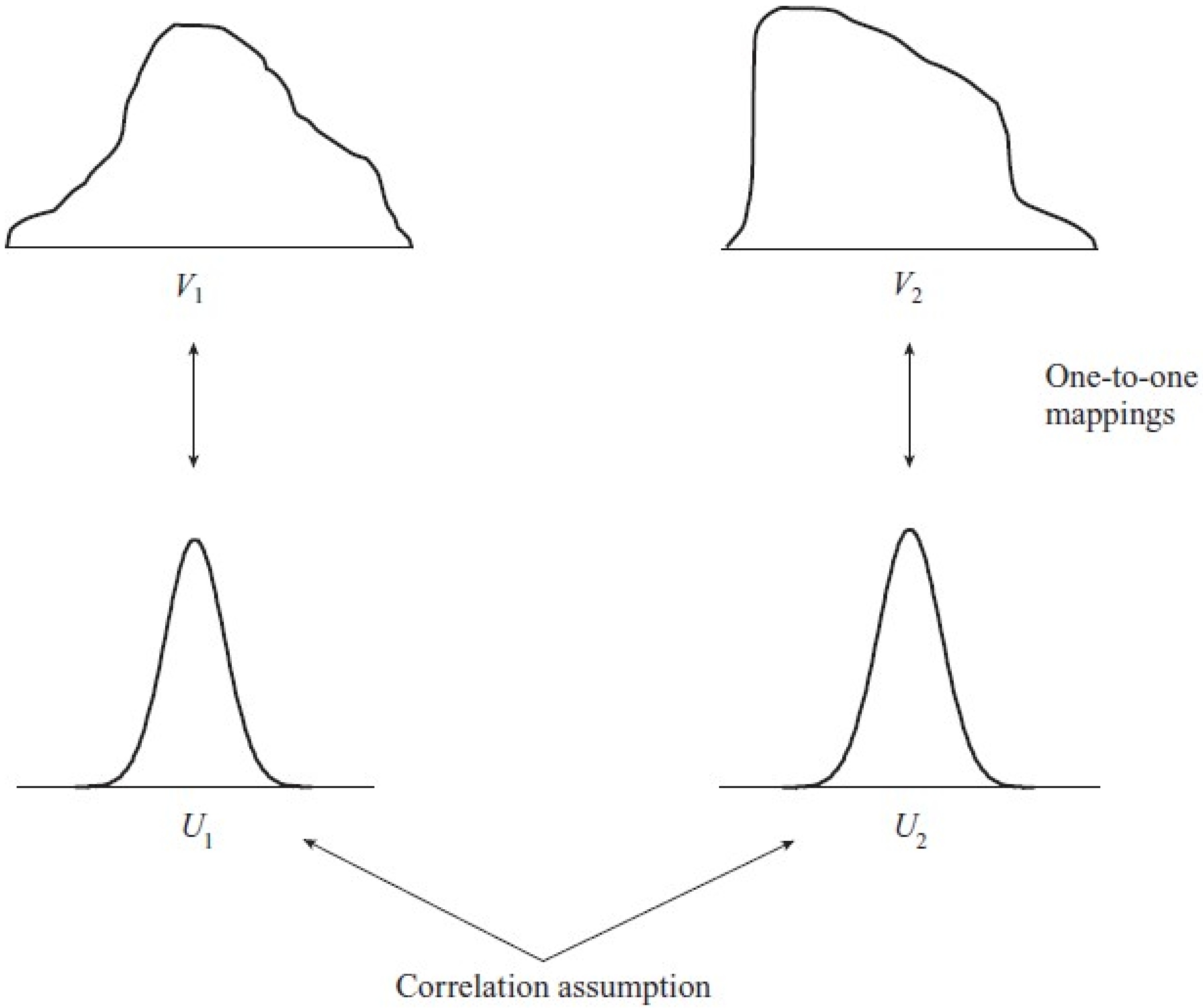
	V_2								
V_1	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	0.006	0.017	0.028	0.037	0.044	0.048	0.049	0.050	0.050
0.2	0.013	0.043	0.081	0.120	0.156	0.181	0.193	0.198	0.200
0.3	0.017	0.061	0.124	0.197	0.273	0.331	0.364	0.381	0.387
0.4	0.019	0.071	0.149	0.248	0.358	0.449	0.505	0.535	0.548
0.5	0.019	0.076	0.164	0.281	0.417	0.537	0.616	0.663	0.683
0.6	0.020	0.078	0.173	0.301	0.456	0.600	0.701	0.763	0.793
0.7	0.020	0.079	0.177	0.312	0.481	0.642	0.760	0.837	0.877
0.8	0.020	0.080	0.179	0.318	0.494	0.667	0.798	0.887	0.936
0.9	0.020	0.080	0.180	0.320	0.499	0.678	0.816	0.913	0.970

(Correlation parameter = 0.5. Table shows the joint probability that V_1 and V_2 are less than the specified values.)

Correlations and Copulas

- **COPULAS:**

The Way in Which a Copula Model Defines a Joint Distribution



Correlations and Copulas

- **TYPES OF COPULAS:**
  - A **Gaussian copula** maps the marginal distribution of each variable to the standard normal distribution. The mapping of each variable to the new distribution is done based on percentiles.
  - The **student's t-copula** is similar to the Gaussian copula. However, variables are mapped to distributions of U_1 and U_2 that have a bivariate Student's t-distribution rather than a normal distribution.

    Procedure to create a Student's t-copula:

    Step 1: Obtain values of λ by sampling from the inverse chi-squared distribution with f degrees of freedom.

    Step 2: Obtain values by sampling from a bivariate normal distribution with correlation ρ.

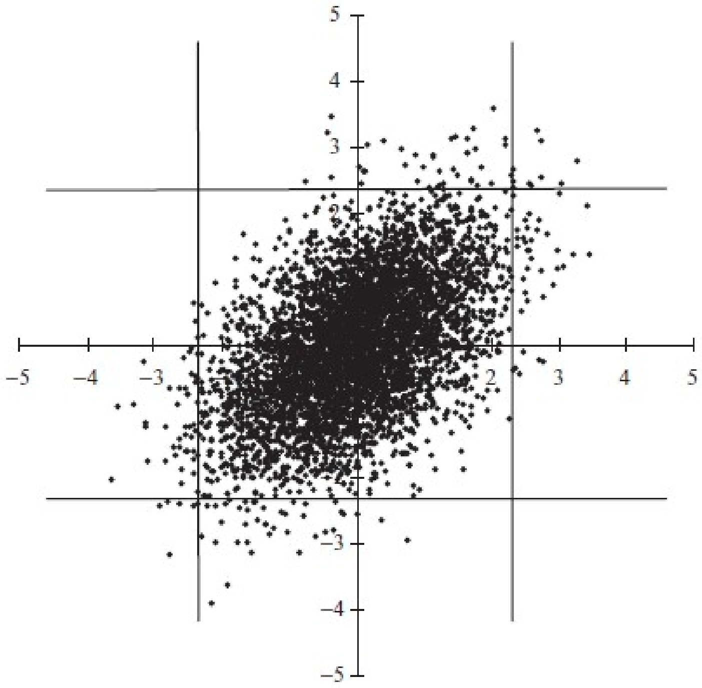
    Step 3: multiply $\sqrt{f/x}$ by the normally distributed samples.
  - A **multi-variance copula** is used to define a correlation structure for more than two variables. Suppose the marginal distributions are known for N variable: V_1 , $V_{2\ldots,}$ $V_{N,.}$ distribution V_i for each i variable is mapped to a standard normal distribution U_i. Thus, the correlation structure for all variable is now based on a multivariate normal distribution.
  - **Factor copula** models are often used to define the correlation structure in multivariate copula models.

$$U_i = a_i F + \sqrt{1 - a_i^2} Z_i$$

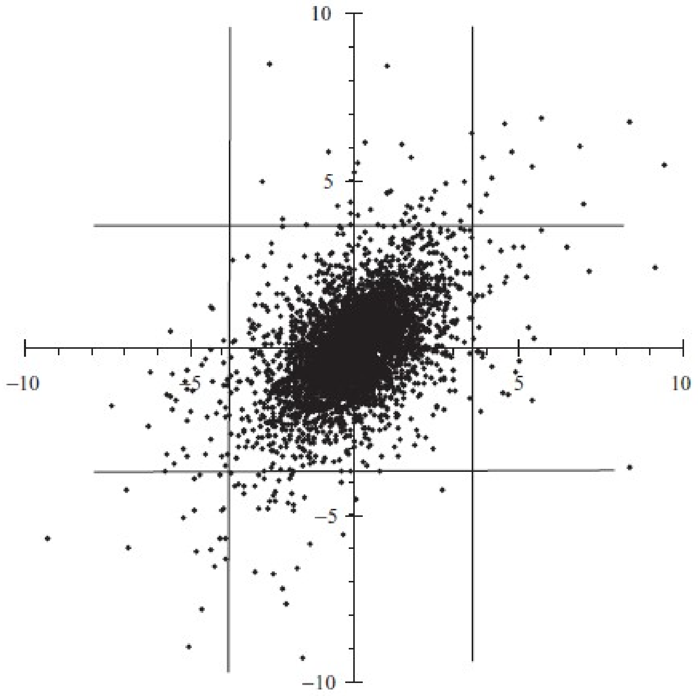
Correlations and Copulas

- **TAIL DEPENDENCE:**
  - There is greater tail dependence in a bivariate Student’s t-distribution than a bivariate normal distribution.
  - During a financial crisis or some other extreme market condition, it is common for assets to be highly correlated and exhibit large losses at the same time.
  - This suggests that the student’s t-copula is better than a Gaussian copula in describing the correlation structure of assets that historically have extreme outliers in the distribution tails at the same time.

5,000 Random Samples from a Bivariate Normal Distribution



5,000 Random Samples from a Bivariate Student *t*-distribution with Four Degrees of Freedom



Thanks

Leading Derivative & Risk Advisors

成蹊风险研究

riskmacro.com