Lecture 6: Multivariate Distributions: Basics, EDA, and Models/Estimation
Statistics 509 – Winter 2022
Reference: Chapters 7 and 8

Brian Thelen
258 West Hall
bjthelen@umich.edu

Outline

- General review of
  - Multivariate distributions
  - Initial measures of dependence
- EDA: Scatter plots and Correlational/Dependency analysis
  - Pearson correlation coefficient and Spearman's rho
- Multivariate Models
  - Multivariate Normal
  - Multivariate t
  - Copulas

Financial data are often of multivariate nature

- Stock price and various explanatory variables (PE, EBIT).
- Stock performance in one time period vs stock performance in previous time period.
- Stocks of same/similar industry
    - Pharmaceutical (Pfizer, Johnson & Johnson, Merck)
    - Automakers (GM and Ford)
    - Google and Microsoft
- Various indices (DJIA, S&P 500)
- Portfolio consisting of stocks and bonds.

Joint Distribution

- Joint CDF

$$F(x_1, \dots, x_p) = P(X_1 \le x_1, \dots, X_p \le x_p)$$

- Joint PDF

$$F(x_1, \dots, x_p) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} f(x_1', \dots, x_p') dx_1' \cdots dx_p'$$

where

$$f(x_1, \dots, x_p) = \frac{\partial^p F(x_1, \dots, x_p)}{\partial x_1 \cdots \partial x_p}$$

Independence

- Marginal density

$$
\begin{aligned}
f_{X_1}(x_1) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n) dx_2 \cdots dx_p \\
&\vdots \\
f_{X_p}(x_p) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_p) dx_1 \cdots dx_{p-1}
\end{aligned}
$$

- Independence

Independent就可以累乘

$$
F(x_1, \ldots, x_p) = F_{X_1}(x_1) \cdots F_{X_p}(x_p)
$$

$$
f(x_1, \ldots, x_p) = f_{X_1}(x_1) \cdots f_{X_p}(x_p)
$$

- Interpretation/Utility of independence

Expectation of Multiple RVs

For random variables $X_1, X_2, \ldots, X_p$ with joint pdf $f(x_1, x_2, \ldots, x_p)$,

- For arbitrary function $h(X_1, X_2, \ldots, X_p)$,

$$E\left[h(X_1, X_2, \ldots, X_p)\right] = \qquad .$$

- For linear combinations $\sum_{j=1}^{p} a_j X_j + b$,

$$E\left[\sum_{j=1}^{p} a_j X_j + b\right] = \sum_{j=1}^{p} a_j E[X_j] + b \qquad .$$

Covariance and Correlation

- With a single variable, we have $E(X)$ and Var(X).
- With two random variables, we would like to measure the joint variation or linear association between them.
- Covariance

$$\begin{aligned} \text{Cov}(X,Y) &= E[(X-E(X))(Y-E(Y))] \\ &= E(XY) - E(X)\cdot E(Y) \end{aligned}$$

- Pearson correlation

$$\text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}$$

  - $-1 \leq \text{Corr}(X,Y) \leq 1$

Measure linear independency

Properties of Covariance

- $\text{Cov}(X, X) = \text{Var}(X)$
- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
- If X and Y are independent, $\text{Cov}(X, Y) = 0$.
- $\text{Cov}(aX, bY) = ab \cdot \text{Cov}(X, Y)$
- Covariance of summation

$$\text{Cov}\left(\sum_{j=1}^{p} a_j X_j, \sum_{j'=1}^{q} b_{j'} Y_{j'}\right) = \sum_{j=1}^{p} \sum_{j'=1}^{q} a_j b_{j'} \text{Cov}(X_j, Y_{j'})$$

- Variance of summation

$$\text{Var}\left(\sum_{j=1}^{p} a_j X_j\right) = \sum a_j^2 \text{Var}(X_j) + \sum_{i \neq j} a_i a_j \text{Cov}(X_i, X_j).$$

- If $X_1, \ldots, X_p$ are independent

$$\text{Var}\left(\sum_{j=1}^{p} a_j X_j\right) = \sum a_j^2 \text{Var}(X_j).$$

Problem Suppose $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables with mean μ and variance σ^2. Derive the mean and variance for

$$S_n = \sum_{i=1}^{n} X_i \quad \text{and} \quad \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Answer. .

$E[S_n] = n\mu$

$VaR[S_n] = n\sigma^2$

$E[\overline{X}_n] = \mu$

$VaR[\overline{X}_n] = \frac{\sigma^2}{n}$

Mean and Covariance of Random Vectors

- If $\mathbf{X}$ is a vector of random variables, $\mathbf{X} = (X_1, \dots, X_p)^\top$,

$$E(\mathbf{X}) \quad = \quad [E(X_1), \dots, E(X_p)]^\top$$

- The covariance matrix of $\mathbf{X}$ is $p \times p$ matrix

$$\mathrm{COV}(\mathbf{X}) = \begin{pmatrix} \mathrm{Var}(X_1) & \cdots & \cdots & \mathrm{Cov}(X_1, X_p) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Var}(X_2) & \cdots & \mathrm{Cov}(X_2, X_p) \\ \vdots & \ddots & \ddots & \vdots \\ \mathrm{Cov}(X_p, X_1) & \cdots & \cdots & \mathrm{Var}(X_p) \end{pmatrix}$$

  - Alternative Formula: $\mathrm{COV}(\mathbf{X}) = E[(\mathbf{X} - E(\mathbf{X}))(\mathbf{X} - E(\mathbf{X}))^\top]$
- For two random vectors $\mathbf{X}$ and $\mathbf{Y}$ (p- and q-dimensional)

$$\mathrm{COV}(\mathbf{X}, \mathbf{Y}) = E\left[(\mathbf{X} - E(\mathbf{X}))(\mathbf{Y} - E(\mathbf{Y}))^\top\right] \qquad (p \times q)$$

Formulas & Properties of Mean/Covariances

- Linear transformations: $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$

$$E(\mathbf{Y}) = \mathbf{A}E(\mathbf{X}) + \mathbf{b}$$

$$\text{COV}(\mathbf{Y}) = A\,Cov(X)A^T$$

都是半正定矩阵

- Covariance matrices are symmetric and non-negative definite, i.e., for covariance matrix $\mathbf{\Sigma} = \text{COV}(\mathbf{X})$ and vector $\mathbf{u}$,

$$\mathbf{u}^\top \mathbf{\Sigma} \mathbf{u} \geq 0 = VaR[u^T X]$$

  - Eigenvalues of covariance matrix are all non-negative

- For $\mathbf{X}$ and $\mathbf{Y}$, with appropriate matrices $\mathbf{A}, \mathbf{B}$ and vectors $\mathbf{u}, \mathbf{v}$,

$$\text{COV}(\mathbf{A}\mathbf{X}, \mathbf{B}\mathbf{Y}) = A\,Cov[X,Y]\,B^T \quad .$$

$$\text{COV}(\mathbf{u}^\top\mathbf{X}, \mathbf{v}^\top\mathbf{Y}) = u^T Cov[X,Y]\,v \quad .$$

Correlation Matrix of Random Vectors

- The correlation matrix $\text{CORR}(\mathbf{X})$ is

$$\begin{pmatrix} 1 & \cdots & \cdots & \text{Corr}(X_1, X_p) \\ \text{Corr}(X_2, X_1) & 1 & \cdots & \text{Corr}(X_2, X_p) \\ \vdots & \ddots & \ddots & \vdots \\ \text{Corr}(X_p, X_1) & \cdots & \cdots & 1 \end{pmatrix}$$

- Diagonal elements being 1 – corresponds to $\text{Corr}(X_i, X_i) = 1$
- Correlation matrices are simply the ... .

Covariance matrix of $X_1', X_2' \cdots X_p'$ after standardizing

$$X_n' = \frac{X_n}{\sigma_n}$$

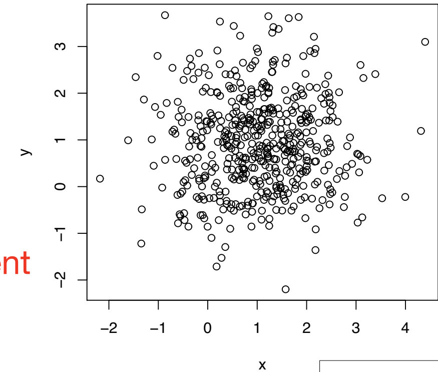
- Thus $\text{CORR}(X)$ is non-negative definite

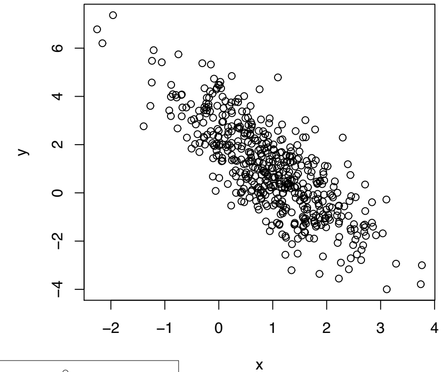
Bivariate Data: Scatter Plots

Remark. For bivariate data $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ it is always very natural to look at scatter plots – allows for seeing

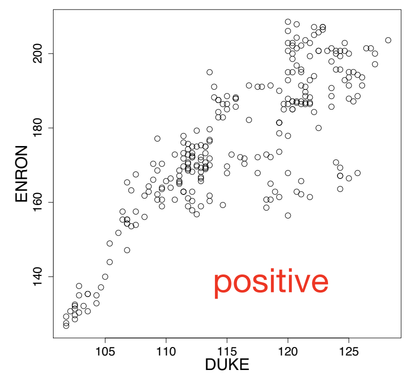
- Spread
- Correlations
- Patterns

independent





negative



R-command for Scatter Plot:

```
plot(x,y,xlab='x',ylab='y')
```

Linear Dependence: Pearson Correlation

- Definition

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\cdot\sqrt{\mathrm{Var}(Y)}}.$$

- Suppose have sample $(x_1, y_1), \ldots, (x_n, y_n)$ and want estimate of correlation ρ
- Natural to use **Pearson correlation coefficient**

$$\hat{\rho} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \cdot \sum_{i=1}^{n}(y_i - \overline{y})^2}}$$

where

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

Remarks.

- Both $-1 \leq \rho \leq 1$ and $-1 \leq \hat{\rho} \leq 1$
- The Pearson correlation coefficient is not so good for heavy-tailed distributions use transformation to shrink the tail
  - One solution – transform the original data so as to remove the effect of heavy tails

Another Measure of Dependence: Spearman's Correlation

- Definition

把变量变成uniform distribution,然后找关系

描述monotonic dependency

$$\rho_S = \rho(F_X(X), F_Y(Y))$$

- Empirical estimate from data $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$
  - replace (X, Y) be discrete rv approximation (X_n, Y_n)
  - replace F_X by $\hat{F}_{Xn}$ and F_Y by $\hat{F}_{Yn}$

  $V = U^2$

  $\rho\left[\frac{U}{b}, F_V(V)\right] = 1$

  The result is

$$\hat{\rho}_S = \frac{12}{n(n^2-1)} \sum_{i=1}^{n} \left(\text{rank}(x_i) - \frac{n+1}{2} \right) \left(\text{rank}(y_i) - \frac{n+1}{2} \right)$$

本质是对rank计算Pearson Correlation

  - $\text{rank}(x_i)$ is rank of x_i in the sample $x_1, x_2, \ldots, x_n$
  - $\text{rank}(y_i)$ is rank of y_i in the sample $y_1, y_2, \ldots, y_n$
- Correlation of ranks: Depends on the relative rankings of x_i and y_i.
- Remove the relative sizes of the values of X among themselves, similarly for Y.

Spearman's Correlation: Properties

Problem. **(a)** Suppose g and h are monotonically increasing functions with X, Y are random variables. Derive an expression for Spearman correlation between $g(X)$ and $h(Y)$.

(b) Suppose g and h are monotonically increasing functions with data $\{x_i, y_i\}_1^n$ are random variables. Derive an expression for Spearman correlation between $\{g(x_i), h(y_i)\}$.

(c) Suppose X, Y are random variables – what can you say about the spearman correlation between e^X and $4Y^3$?

(d) What happens in **(c)** for the Pearson correlation?

Answer. .

(a): 和X,Y相同

(b): g和h都是单调递增,因此rank不变,因此spearman correlation不变

(c): 和X,Y相同

(d): 会改变

Answer.

Coffee Data

- Daily closing prices of Brazilian and Colombian coffee and the corresponding log-returns.
- Remove the zeros.

```
> data(BCofLRet)
> data(CCofLRet)
> NZ <- (BCofLRet != 0 & CCofLRet !=0)
> BLRet <- BCofLRet[NZ]
> CLRet <- CCofLRet[NZ]
```

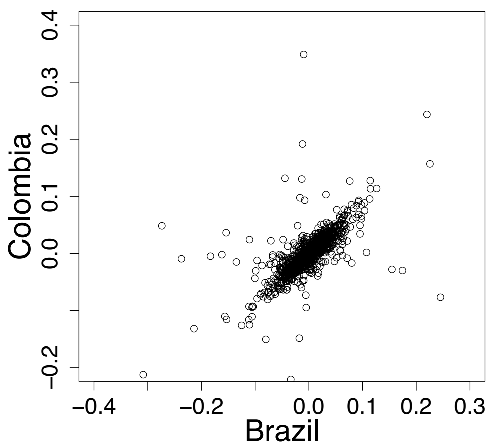
Coffee vs Normal Distribution

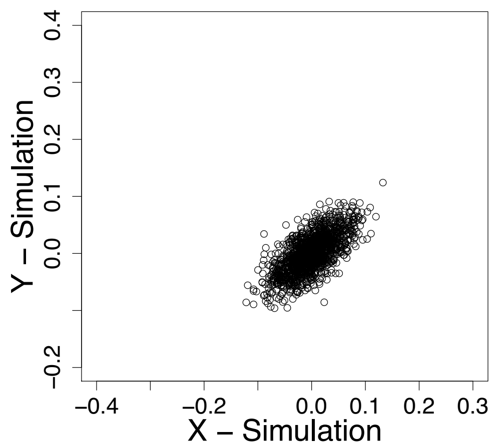
```
> myxlim <- c(-0.4, 0.3)
> myylim <- c(-0.2, 0.4)
> plot(BLRet, CLRet, xlim=myxlim, ylim=myylim,
    xlab="Brazil", ylab="Colombia")

> mu <- c(mean(BLRet), mean(CLRet))
> Sigma <- var(cbind(BLRet, CLRet))
> n <- length(BLRet)
> nsim <- rmvnorm(n, mean=mu, cov=Sigma)
> plot(nsim, xlim=myxlim, ylim=myylim,
    xlab="X - Simulation", ylab="Y - Simulation")
```

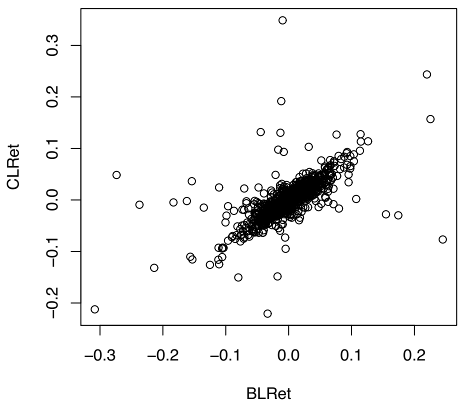
和模拟结果差很多,因此不是双变量正态分布

但是可能是X^3和Y^3的分布(X,Y服从双变量正态分布)





Coffee Data – Brazil and Columbian coffee futures



$$\Sigma\mathrm{cov} = \frac{1}{n-1}\sum_{i}^{n}(x-\bar{x})(x-\bar{x})^T$$

$$\bar{x} = \frac{1}{n}\sum x$$

R-commands Pearson, Spearman, and Kendall correlations

```
> cor(BLRet,CLRet)  # Default for cor is Pearson.
[1] 0.6901319
>
> cor(BLRet,CLRet, method = "spearman")
[1] 0.8356541
```

Spearman更好,因为Pearson容易受到outlier影响

CRSP data: Analysis in R

- Data from R's Ecdat: GE, IBM, and Mobil along with CRSP value-weighted index
- Daily returns from Jan 3 1989 to Dec 31 1998

```
> library(Ecdat)
> data(CRSPday)
> CRSPday[c(1:4),]
     year month day        ge       ibm     mobil      crsp
[1,] 1989     1   3 -0.016760  0.000000 -0.002747 -0.007619
[2,] 1989     1   4  0.017045  0.005128  0.005510  0.013016
[3,] 1989     1   5 -0.002793 -0.002041  0.005479  0.002815
[4,] 1989     1   6  0.000000 -0.006135  0.002725  0.003064

> signif(cov(CRSPday[,c(4:7)]),digits=3)
            ge      ibm    mobil     crsp
ge    1.88e-04 8.01e-05 5.27e-05 7.61e-05
ibm   8.01e-05 3.06e-04 3.59e-05 6.60e-05
mobil 5.27e-05 3.59e-05 1.67e-04 4.31e-05
crsp  7.61e-05 6.60e-05 4.31e-05 6.02e-05
```

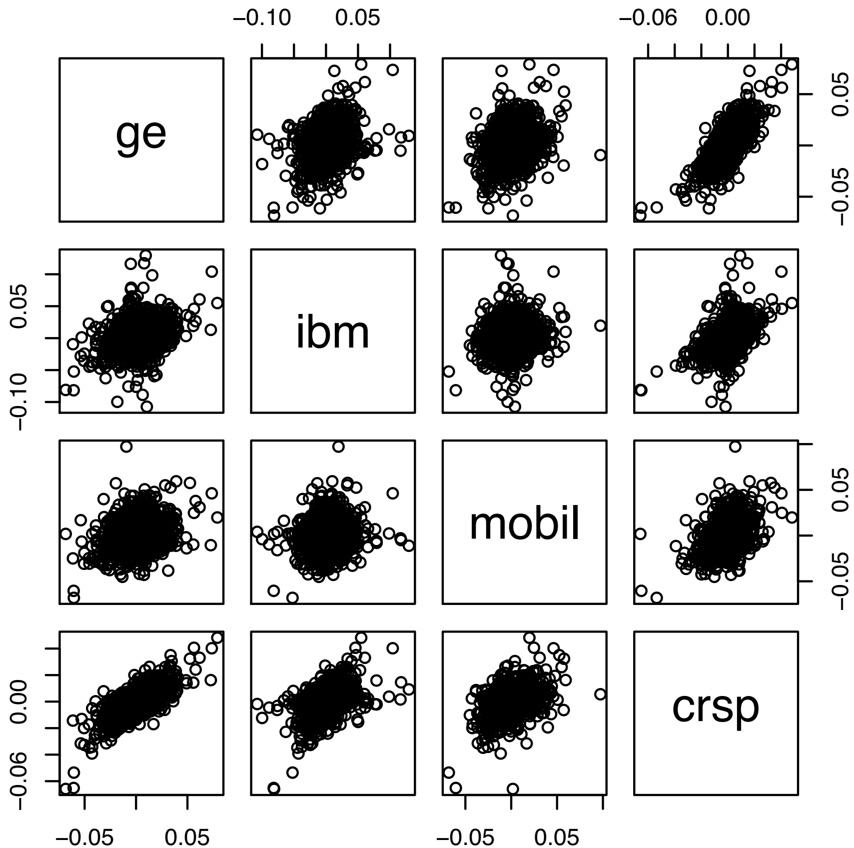
CRSP data: Correlational Analysis in R

```
> signif(cor(CRSPday[,c(4:7)]),digits=3)
         ge   ibm mobil  crsp
ge    1.000 0.334 0.297 0.715
ibm   0.334 1.000 0.159 0.486
mobil 0.297 0.159 1.000 0.429
crsp  0.715 0.486 0.429 1.000

> signif(cor(CRSPday[,c(4:7)],method="spearman"),digits=3)
         ge   ibm mobil  crsp
ge    1.000 0.315 0.272 0.638
ibm   0.315 1.000 0.165 0.474
mobil 0.272 0.165 1.000 0.399
crsp  0.638 0.474 0.399 1.000

> pairs(CRSPday[,c(4:7)])
```

画出所有变量之间的散点图



ge
ibm
mobil
crsp
−0.10
0.05
−0.06
0.00
0.05
−0.05
0.05
−0.10
0.05
−0.05
0.00
−0.06
−0.05
0.05
−0.05
0.05

Multivariate Normal Distribution

- $\mathbf{X} = [X_1, \ldots, X_p]^\top$ is said to be multivariate normal if it has joint density

$$f_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(x_1, \ldots, x_p) = \frac{1}{\sqrt{(2\pi)^p \cdot \det(\boldsymbol{\Sigma})}} \exp^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

  where $E(\mathbf{X}) = \boldsymbol{\mu}$ and covariance matrix of $\mathbf{X}$ is $\boldsymbol{\Sigma}$.
- Notation: $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
- Widespread use in portfolio theory. Not so widely now

Multivariate Normal Distribution: Properties

- For $\mathbf{X} \sim \mathcal{N}(\mu, \mathbf{\Sigma})$, have that $X_1, \ldots, X_p$ are independent if and only if $\mathbf{\Sigma}$ is a 对角矩阵 . In this case,

$$f_{\boldsymbol{\mu},\mathbf{\Sigma}}(x_1, \ldots, x_p) = \prod f_{\mu_i \sigma_i}(x)$$

- Linear transformation of X:

  If $\mathbf{X} \sim N(\boldsymbol{\mu}, \mathbf{\Sigma})$ and $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ where $\mathbf{A}$ is a $q \times p$ matrix, with $q \leq p$ and $\mathbf{b}$ is a q-dimensional vector, then

$$\mathbf{Y} \sim N(A\mu+b, A\Sigma A^T) .$$

  - Special case of $\mathbf{u}^\top \mathbf{X} + b$ when $\mathbf{u}$ is $d \times 1$ vector and b is a number – in this case,

$$\mathbf{Y} \sim N(u^T\mu+b, u\Sigma u^T) \quad \text{变为单变量} .$$

Simple Risk Management Example: Portfolio with 2 assets

- Initial value

$$V_0 = n_1 P_1 + n_2 P_2$$

- New/future value

$$V = n_1 P_1' + n_2 P_2'$$

- Returns/weights

$$R_1 = \frac{P_1'}{P_1}, \qquad R_2 = \frac{P_2'}{P_2}$$

Proportion in asset 2

Proportion in asset 1

$$w_1 = \frac{n_1 P_1}{n_1 P_1 + n_2 P_2}, \qquad w_2 = \frac{n_2 P_2}{n_1 P_1 + n_2 P_2}.$$

- Portfolio Return

log return can't work anymore
需要研究r1和r2的联合分布

$$R = \frac{n_1 P_1' + n_2 P_2'}{n_1 P_1 + n_2 P_2} = \quad | = w_1 R_1 + w_2 R_2$$

VaR_q and ES

Assumption:

$$R = w_1 R_1 + w_2 R_2 = w_1 R_1 + (1-w_1) R_2$$

如果R1,R2服从不相关的双正态分布

$$R \sim N(w_1\mu_1 + (1-w_1)\mu_2,\ w_1^2\sigma_1^2 + 2w_1(1-w_1)Cov[R_1,R_2] + (1-w_1)^2\sigma_2^2$$

- Value at risk. Need to solve

$$q = P(R \leq -r)$$

- Expected shortfall. Need to compute

$$E(-R \mid -R > VaR_q) = \frac{1}{q}\int_{-\infty}^{-VaR} -rF_R(dr)$$

Remark. Need the distribution of R – can derive from

Example. Suppose have two assets and under some model, the returns R_1, R_2 are bivariate normal.

(a) Suppose $E(R_1) = E(R_2) = .02, \text{SD}(R_1) = .015, \text{SD}(R_2) = .025, \text{Corr}(R_1, R_2) = 0$. Derive the optimal allocation between the assets, i.e., optimal relative weights w_1 and w_2.

(b) Repeat **(a)** , with $\text{Corr}(R_1, R_2) = .6$.

(c) Repeat **(a)** with $E(R_1) = .015$, $\text{Corr}(R_1, R_2) = .6$ giving specifics for a methodology finding the optimal allocation between the two assets, relative to minimizing $\widetilde{\text{VaR}}$ at $\alpha = .005$.

Answer.

(a): $w \in [0,1]$, $R_w = wR_1 + (1-w)R_2$

$E[R] = 0.02$, 因此要最小化 variance

$\sigma^2 = w^2 \cdot (0.015)^2 + (1-w)^2 (0.025)^2$

$\frac{d\sigma^2}{dw} = 2w(0.015)^2 - 2(1-w)(0.025)^2 = 0$

$w \approx 0.75$ 鼓励资产分散投资不相关资产

Answer.

(b) $E[R] = 0.02$

$$\sigma^2 = w^2 (0.015)^2 + (1-w)(0.025)^2 + 2w(1-w)\, 0.015 \cdot 0.025$$

$$\frac{d\sigma^2}{dw} = 0 \Rightarrow w \approx 0.87$$

相关性越强,投资越不分散

$$W = \frac{\sigma_2^2 - \rho_{12}\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2}$$

Answer.

(c):

1. 使用不同比例的w生成portfolio收益率
2. 计算收益率的均值和方差
3. 计算VaR,并选取最小VaR对应的w

Answer

Multivariate t-Distribution

$w^T Y$服从$t(w^T\mu, w^T\Sigma w)$的t分布

- $\mathbf{Y} = [Y_1, \ldots, Y_p]^\top$ is said to have multivariate $t_\nu(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ distribution if $\Sigma = \frac{\nu}{\nu-2}\Lambda$

$$\mathbf{Y} \sim \boldsymbol{\mu} + \sqrt{\frac{\nu}{W}}\mathbf{X} = \mu + \frac{X}{\sqrt{\frac{W}{\nu}}}$$

  where
  - $\mathbf{X} \sim \mathcal{N}(0, \boldsymbol{\Lambda})$ X的协方差矩阵
  - W is chi-square rv with ν degrees of freedom
  - $\mathbf{X}$ and W are independent
- Often refer to $t_\nu(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ as a scale mixture of multivariate normal
- Note that W multiplies all of components at the same time - implies "tail dependence"
- Note that tails of each marginal distribution have same ν - implications tail都相同
- Is used in portfolio theory.

Plots, Mean and Covariance of t-distribution

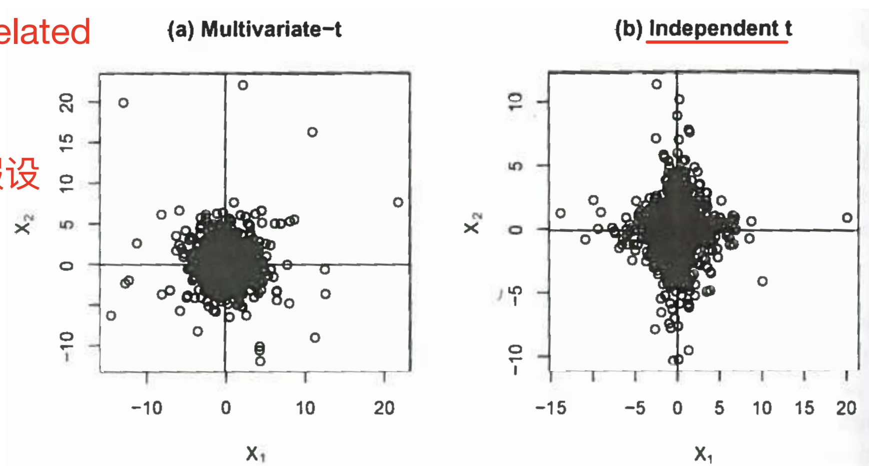
- Suppose $\mathbf{X} \sim t_\nu(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ - then
  - $E(\mathbf{X})$ exists if $\nu >$ 1 and is equal to μ
  - Cov$(\mathbf{X})$ exists if $\nu >$ 2 and is equal to $\frac{\nu}{\nu-2}\Lambda$

- Interpetation when Λ is diagonal matrix
  - Uncorrelated vs. independence
  - Comparison of bivariate t with independent t

双变量T分布
each marginal is T

each marginal uncorrelated
但是dependent

只有dof很大时,可以假设
independence

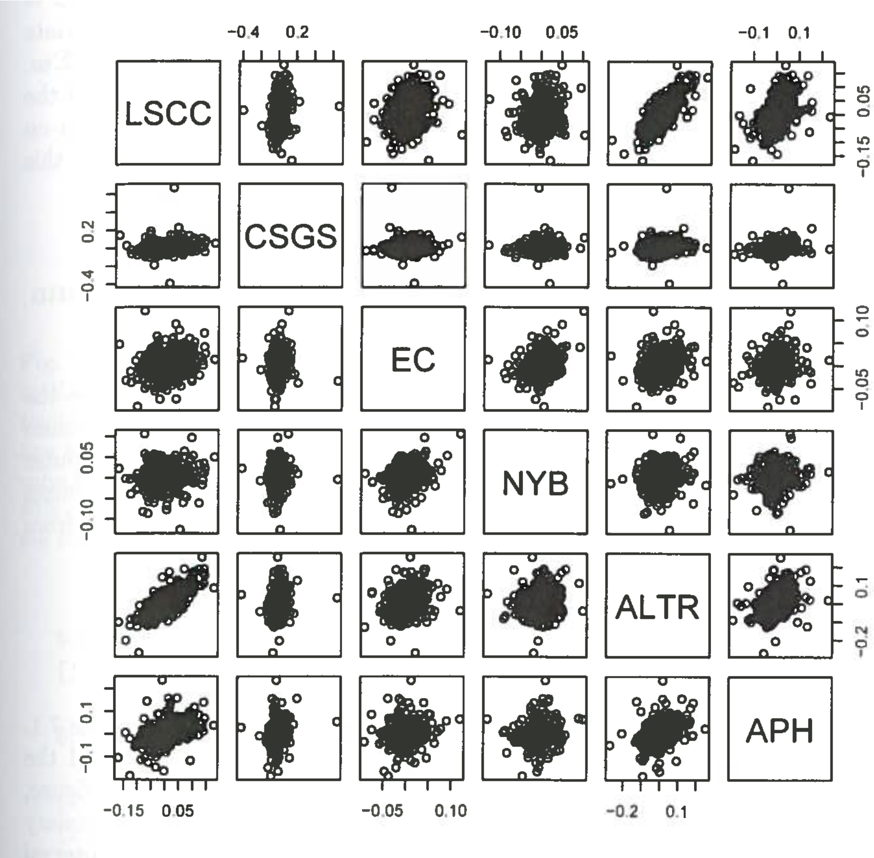


Implications

两边都是uncorrelated,但是球状的是dependent

Multivariate t-distribution: tail dependencies

- Bivariate plots of 6 Midcap stocks returns from fEcofin (cf. Ruppert)



Interpretation

Problem. Suppose $\mathbf{X} \sim t_\nu(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ and $\mathbf{A}$ is a $q \times p$ matrix, with $q \leq p$ and $\mathbf{b}$ is a q-dimensional.

(a) What is the distribution of $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$?

(b) What can you say about the distribution of random variables $X_1, X_2, \ldots, X_p$?

(c) What can you say about the distribution of random vector $\mathbf{X}$ when ν is large?

(d) What can you say about the dependence/independence between the random variables $X_1, X_2, \ldots, X_p$.

Answer. . (a) $Y = AX + b = \underbrace{A\mu + b}_{\mu} + \frac{AZ}{\sqrt{\frac{W}{\nu}}} \sim t_\nu(A\mu + b, A\Lambda A^T)$

(b) $X_i \sim t_\nu(\mu_i, \Lambda_{ii})$

(c) Close to multivariant normal distribution

(d) All dependent, $\sqrt{\frac{\nu}{W}}$ is common to all

Multivariate Normal: Parameter Estimation

- For p-dimensional data (think returns of p assets), $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, the MLE for mean and covariance of multivariate normal distributional model is

MLE的估计值就是样本的值

sample mean $\rightarrow$ $\hat{\boldsymbol{\mu}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i = \overline{\mathbf{x}}$

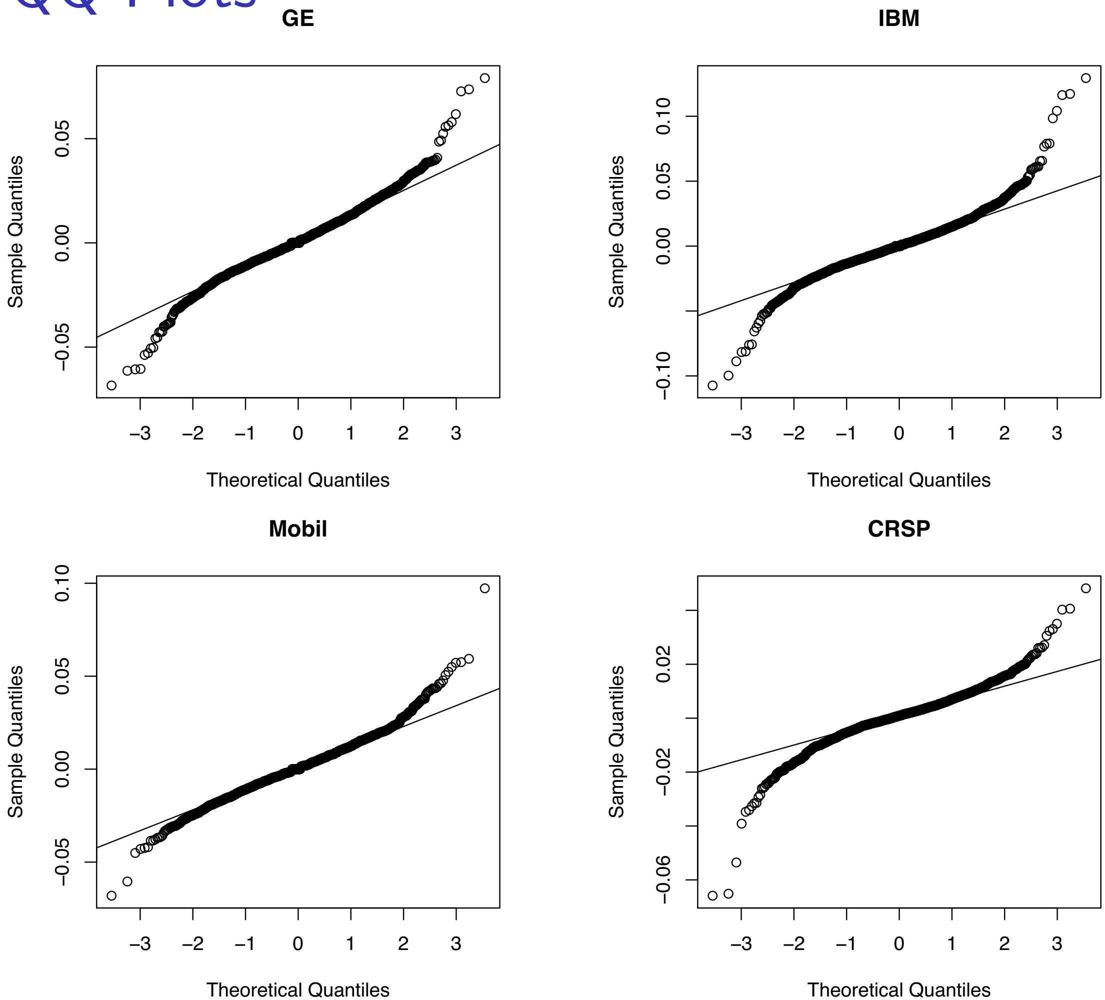
sample covariance matrix $\rightarrow$ $\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n} (\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^\top$

R-commands: MLE estimates for multivariate normal

```
> library(Ecdat)
> data(CRSPday)
> CRSPday[c(1:4),]
     year month day        ge       ibm     mobil      crsp
[1,] 1989     1   3 -0.016760  0.000000 -0.002747 -0.007619
[2,] 1989     1   4  0.017045  0.005128  0.005510  0.013016
[3,] 1989     1   5 -0.002793 -0.002041  0.005479  0.002815
[4,] 1989     1   6  0.000000 -0.006135  0.002725  0.003064

> library(timeSeries)
> sampmean = signif(colMeans(CRSPday[,c(4:7)]),digits=3)
> sampmean
      ge      ibm    mobil     crsp
0.001070 0.000700 0.000779 0.000678
> signif(cov(CRSPday[,c(4:7)]),digits=3)
            ge      ibm    mobil     crsp
ge    1.88e-04 8.01e-05 5.27e-05 7.61e-05
ibm   8.01e-05 3.06e-04 3.59e-05 6.60e-05
mobil 5.27e-05 3.59e-05 1.67e-04 4.31e-05
crsp  7.61e-05 6.60e-05 4.31e-05 6.02e-05
```

Normal QQ Plots



Implications:

Multivariate t-distribution: MLE

- Want to use MLE in estimating $\boldsymbol{\mu}, \ \boldsymbol{\Lambda}, \nu$
  - Profile likelihood method - for fixed ν, can generate MLE of $\boldsymbol{\mu}$ and $\boldsymbol{\Lambda}$, i.e., then optimize over ν
  - Full MLE (doing in homework)
- Would like to determine if the estimated t-distribution is an appropriate model
  - Use univariate QQ plots
  - Use model-selection criteria - AIC and BIC

Profile Likelihood Estimation

- General problem: MLE of $\boldsymbol{\theta} = (\theta_1, \boldsymbol{\theta}_2)$ with model $f(\mathbf{x}; \boldsymbol{\theta})$ and data $\mathbf{x}_1, \ldots, \mathbf{x}_n$
  - e.g., multivariate t with $\theta_1 = \nu$, and $\boldsymbol{\theta}_2 = (\boldsymbol{\mu}, \boldsymbol{\Lambda})$ and $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are vectors are returns on suite of assets at times $1, 2, \ldots, n$
- Standard MLE: MLE estimation of multiple parameters, need to maximize log-likelihood

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \ \ell(\boldsymbol{\theta}) = \arg\max_{\theta_1, \boldsymbol{\theta}_2} \ \ell(\theta_1, \boldsymbol{\theta}_2)$$

- Profile likelihood (alternative): Solve for maximizing values in a "nested" fashion

Step 1 : $\ell_{max}(\theta_1) = \max_{\boldsymbol{\theta}_2} \ell(\theta_1, \boldsymbol{\theta}_2)$

适合获得confidence interval

Step 2 : $\hat{\theta}_1 = \arg\max_{\theta_1} \ \ell_{max}(\theta_1)$

Step 3 : $\hat{\boldsymbol{\theta}}_2 = \arg\max_{\boldsymbol{\theta}_2} \ \ell(\hat{\theta}_1, \boldsymbol{\theta}_2)$

Profile Likelihood Estimation: Confidence Intervals

- Can construct confidence intervals for profile likelihood estimators
  - Utilizes theory of generalized likelihood ratio tests
- (Relevant) Generalized Likelihood Ratio Tests

$$H_o : \theta_1 = \theta_{0,1} \quad \text{vs.} \quad H_1 : \theta_1 \neq \theta_{0,1}$$

  - Under H_o (and large n),

$$2(\ell_{max}(\hat{\theta}_1) - \ell_{max}(\theta_{0,1})) \sim \chi^2_1$$

  - Decision Rule: Reject H_o if

$$2(\ell_{max}(\hat{\theta}_1) - \ell_{max}(\theta_{0,1})) > \chi^2_{\alpha,1}$$

  where $(1-\alpha)$ quantile

- Generate $100(1-\alpha)\%$ confidence interval with

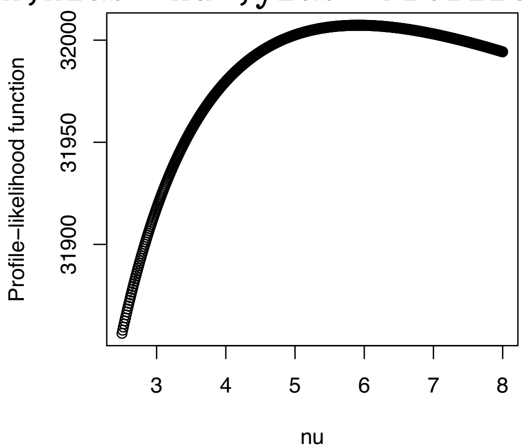
$$100(1-\alpha)\%\ \text{CI} = \theta_1 : l_{max}(\theta) \geq l_{max}(\hat{\theta}) - \frac{1}{2}\chi^2_{\alpha} \quad .$$

Example: Profile likelihood estimation for multivariate t

- Can use profile estimation for multivariate t
- Natural to follow up with QQ plots to check

Example with R-code

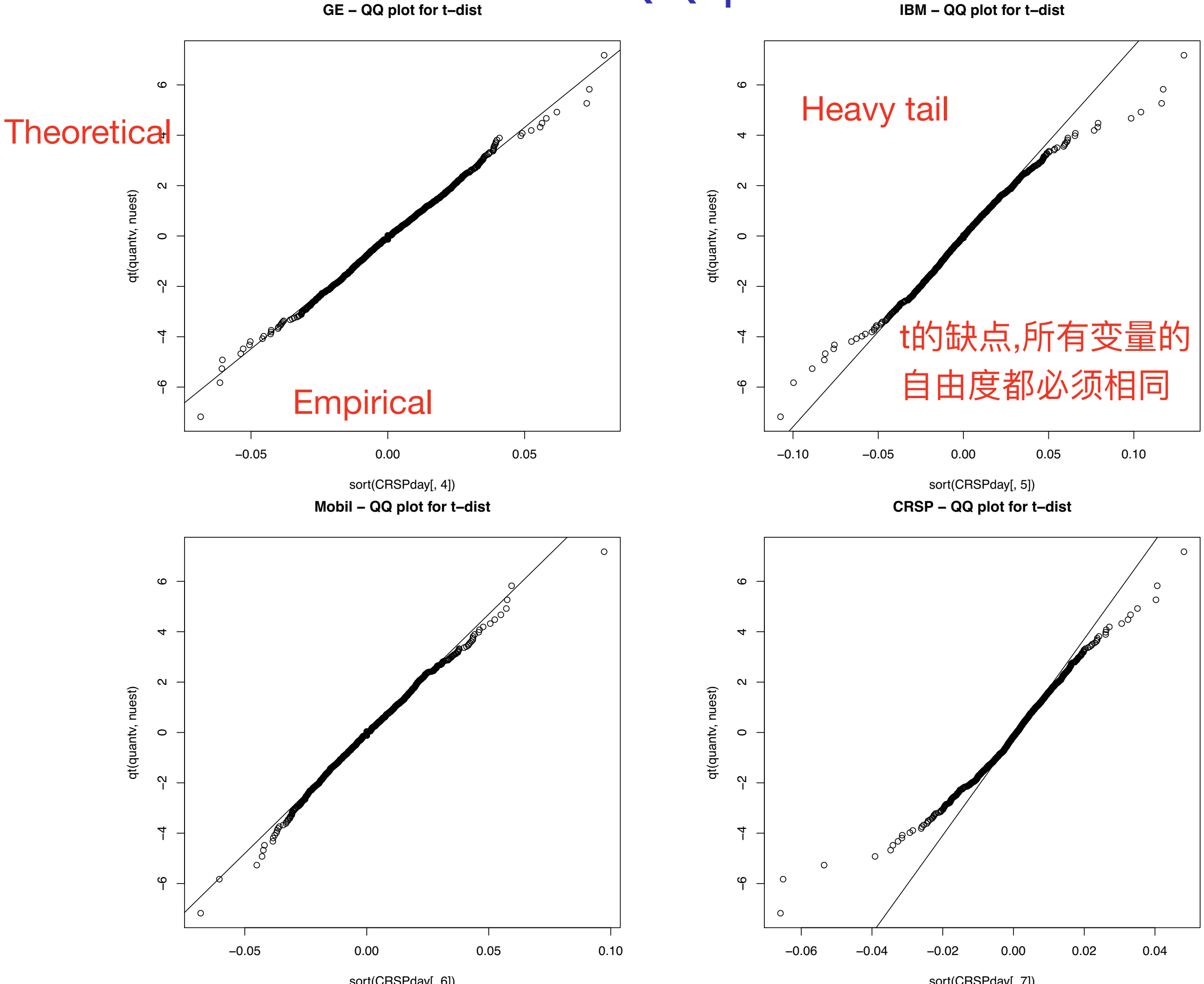
```
> library(mnormt) # needed for dmt
> df = seq(2.5,8,.01)
> n = length(df)
> loglik_max = rep(0,n)
> for(i in 1:n){
+ fit = cov.trob(CRSPday[,c(4:7)],nu=df[i])
+ mu = as.vector(fit$center)
+ sigma =matrix(fit$cov,nrow=4)
+ loglik_max[i] = sum(log(dmt(CRSPday[,c(4:7)],mean=fit$center,
+ S=fit$cov,df=df[i])))}
> plot(df,loglik_max,xlab='nu',ylab='Profile-likelihood function')
```



R-code/results continuation

```
> nuest = df[which.max(loglik_max)]
> nuest
[1] 5.91
> fitfinal = cov.trob(CRSPday[,c(4:7)],nu=nuest)
> fitfinal$center
          ge          ibm        mobil         crsp
0.0009405691 0.0004468076 0.0006876304 0.0007693982
> fitfinal$cov
                ge          ibm        mobil         crsp
ge    1.253813e-04 4.822132e-05 3.422277e-05 4.582990e-05
ibm   4.822132e-05 1.820906e-04 2.297246e-05 3.835559e-05
mobil 3.422277e-05 2.297246e-05 1.154352e-04 2.796438e-05
crsp  4.582990e-05 3.835559e-05 2.796438e-05 3.662112e-05
```

CRSP - multivariate t and QQ plots



Implications:

R-commands

```
> N = length(CRSPday[,4])
> quantv = (1/N)*seq(.5,N-.5,1)
> qqplot(sort(CRSPday[,4]),qt(quantv,nuest),main='GE - QQ plot for t-dist')
> abline(lm(qt(c(.25,.75),nuest)~quantile(CRSPday[,4],c(.25,.75))))

> qqplot(sort(CRSPday[,5]),qt(quantv,nuest),main='IBM - QQ plot for t-dist')
> abline(lm(qt(c(.25,.75),nuest)~quantile(CRSPday[,5],c(.25,.75))))

> qqplot(sort(CRSPday[,6]),qt(quantv,nuest),main='Mobil - QQ plot for t-dist')
> abline(lm(qt(c(.25,.75),nuest)~quantile(CRSPday[,6],c(.25,.75))))

> qqplot(sort(CRSPday[,7]),qt(quantv,nuest),main='CRSP - QQ plot for t-dist')
> abline(lm(qt(c(.25,.75),nuest)~quantile(CRSPday[,7],c(.25,.75))))
```

Beyond Multivariate Normal Distribution

- Modeling dependencies between random variables important in financial applications.
- $f_{X,Y}(x,y)$ vs $f_X(x)f_Y(y)$
- Remove the dependence on marginal distributions; focus on the intrinsic dependence between variables.

Copulas

- Definition: Joint distribution of uniformly distributed random variables. If $U_1, \ldots, U_p$ are $U(0,1)$, then

$$C(u_1, \ldots, u_p) = P(U_1 \leq u_1, \ldots, U_p \leq u_p)$$

is a copula. 把随机变量转化为uniform distribution

- Important fact: X random variable with CDF $F(\cdot) \Rightarrow F(X)$ uniform on $[0,1]$.
- If $X_1, \ldots X_p$ are random variables with CDF's $F_{X_1}, \ldots, F_{X_p}$, then the copula of

$$U_1 = F_{X_1}(X_1), \ldots, \underline{U_p = F_{X_p}(X_p)}$$

is called the copula of $(X_1, \ldots, X_p)$.

Properties

- C is unique if $F_{(X_1,\ldots,X_p)}(x_1,\ldots,x_p)$ is continuous.
- $C(u_1,\ldots,u_p)$ is non-decreasing in each u_j.
- $C(1,\ldots,1,u_j,1,\ldots,1) = u_j$.
- C does not change if one replaces any of the X_j's by a non-decreasing function.
  思考rank和spearson correlation
  - Suppose (X,Y) are two random variables and $(V,W) = (X+2, e^Y)$. The only difference between (X,Y) and (V,W) is a known transformation, and so it is natural to say that "dependency" between the two is in some sense the same.
- The joint CDF can be recovered from the copula and the marginal CDF's via

通过多变量CDF获得copula

$$F_{(X_1,\ldots,X_p)}(x_1,\ldots,x_p) = C(F_{X_1}(x_1),\ldots,F_{X_p}(x_p))$$

Example (a) If $F_{\mathbf{X}}$ is the joint cdf of $X_1, X_2, \ldots, X_p$, and C is the copula cdf, derive the formula that

$$F_{\mathbf{X}}(x_1, \ldots, x_p) = C(F_{X_1}(x_1), \ldots, F_{X_p}(x_p))$$

(b) Suppose $X_1, X_2, \ldots, X_p$ are independent – derive the copula model.

(c) If $Y_1, \ldots, Y_n$ has a copula model as given in part **(b)**, are the random variables independent?

(d) Suppose $X_1, X_2, \ldots, X_p$ have a perfect (perhaps nonlinear) positive dependency. Derive the copula model.

Answer. . (a) $F_x(x_1, x_2 \ldots x_p) = P[X_1 \le x_1 \ldots X_p \le x_p] = P[F(X_1) \le F(x_1) \ldots F(X_p) \le F(x_p)]$

$= C[F_{X_1}(x_1), \ldots F_{X_p}(x_p)]$

(b) $C(u_1, \ldots u_p) = \prod_{i=1}^{p} u_i$

(c) Yes

(d) $C = \min\{u_1 \cdots u_p\}$

General Bivariate Distributions

- The joint density $f_{(X,Y)}$ of X and Y can be written as

$$
\begin{aligned}
f(x,y) &= \frac{\partial^2}{\partial x \partial y} F(x,y) = C[F_X(x), F_Y(y)] \\
&= c(F_X(x), F_Y(y)) f_X(x) f_Y(y)
\end{aligned}
$$

  where $c(u,v) = \frac{\partial^2}{\partial u \partial v} C(u,v)$.
- Specify a bivariate distribution by specifying the marginal distributions and a copula. 多变量联合pdf

$$f(x_1, x_2 \dots) = c(F_{X_1}(x_1), F_{X_2}(x_2) \dots) \prod f_{X_i}(x_i)$$

Copula Models

- Independent copula

$$C(u_1, \ldots, u_p) = u_1 \cdots u_p$$

- Gaussian copula $C(u_1, u_2)$ 只有correlation作为参数

$$\frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} e^{-(s^2-2\rho st+t^2)/2(1-\rho^2)} ds dt$$

  - Marginals need not be Gaussian, e.g. Cauchy, t, and so on
  - When marginals are Gaussian, ρ is the Pearson correlation.
- t copula – parameters ν and scale matrix $\boldsymbol{\Omega}$ correlation matrix
  - Correponds to copula of multivariate t distributions
  - We can remove $\boldsymbol{\mu}$ and reduce to a correlation matrix because.....

Copula Models: Overview

- Multivariate data modeling :

Univariate model/estimation for marginals

$\hat{F}_1, \ldots, \hat{F}_p$

$\downarrow$

transform marginals $\hat{F}_1(x_{11}), \hat{F}_1(x_{12}) \ldots$
$\hat{F}_2(x_{21}), \hat{F}_2(x_{22}) \ldots$

$\downarrow$

Copula modeling/estimation

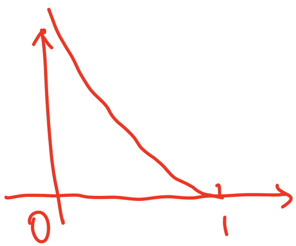
- Have Gaussian-copula and t-copula
- For additional models, will present Archimedean-type (coming)
- Estimation can be
  - Full maximum-likelihood (joint estimation over whole model)
  - Pseudo maximum-likelihood (sequential estimation approach)

Archimedean-type Copulas

$$C(u_1, u_2, \ldots, u_d) = \phi^{-1}(\phi(u_1) + \cdots + \phi(u_d))$$

- ϕ is function on $[0,1]$ and is called the generator – it satisfies that
  - ϕ is a strictly decreasing convex function
  - $\phi(0) = \infty, \ \phi(1) = 0$

Picture



- All of these models, regardless of dimension are examples of exchangeable models - implications are that
- Focus on Gumbel, Clayton, and Frank

dependence between everybody is same

Gumbel (logistic) Copula

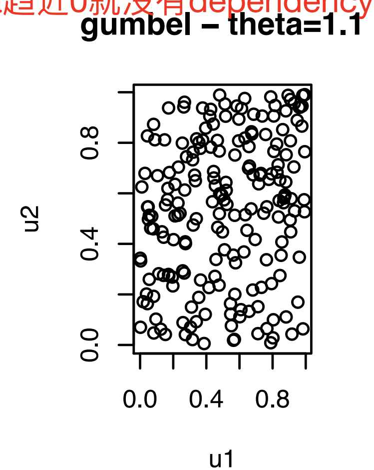
- General d-dimensional Gumbel copula cdf: $\phi(u) = -\log u$

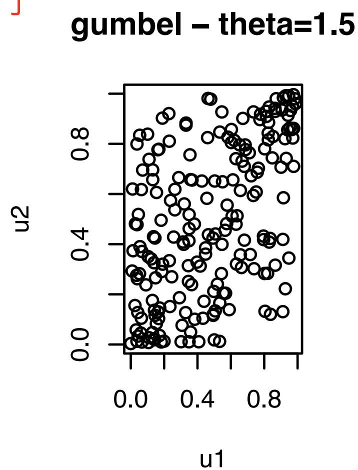
$$C(u_1, \dots, u_d) = e^{-[(-\log u_1)^\theta + \cdots (-\log u_d)^\theta]^{\frac{1}{\theta}}}, \quad \theta \geq 1$$

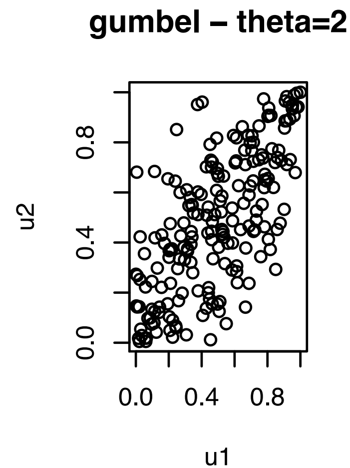
- Independence copula when $\underline{\theta = 1}$
- Copula of random variables with heavy tail distributions.

Gumbel Scatter Plots

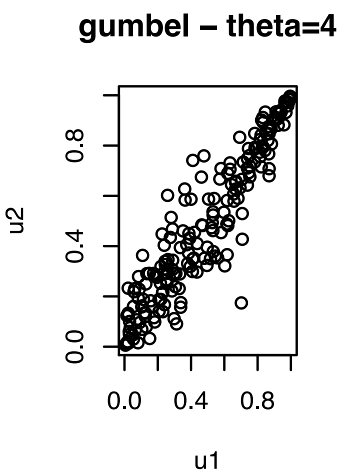
theta趋近0就没有dependency了

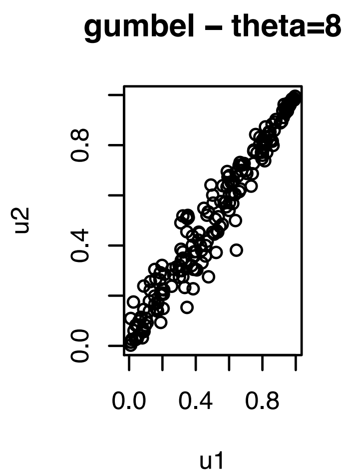


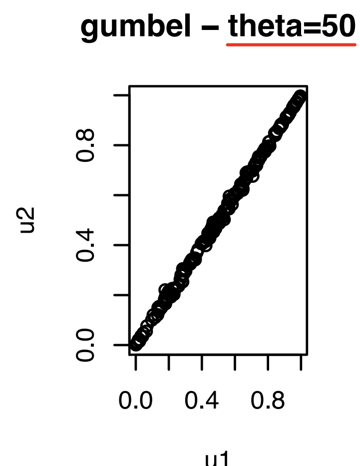




asymmetric model







highly correlated

```
> library(copula)
> par(mfrow=c(2,3))
> c <- gumbelCopula(1.1, dim = 2)
> uv <- rCopula(200, c)
> plot(uv[,1],uv[,2],xlab='u1',ylab='u2',main = 'gumbel - theta=1.1')
> c <- gumbelCopula(1.5, dim = 2)
> uv <- rCopula(200, c)
> plot(uv[,1],uv[,2],xlab='u1',ylab='u2',main = 'gumbel - theta=1.5')
> c <- gumbelCopula(2, dim = 2)
> uv <- rCopula(200, c)
> plot(uv[,1],uv[,2],xlab='u1',ylab='u2',main = 'gumbel - theta=2')
> c <- gumbelCopula(4, dim = 2)
> uv <- rCopula(200, c)
> plot(uv[,1],uv[,2],xlab='u1',ylab='u2',main = 'gumbel - theta=4')
> c <- gumbelCopula(8, dim = 2)
> uv <- rCopula(200, c)
> plot(uv[,1],uv[,2],xlab='u1',ylab='u2',main = 'gumbel - theta=8')
> c <- gumbelCopula(50, dim = 2)
> uv <- rCopula(200, c)
> plot(uv[,1],uv[,2],xlab='u1',ylab='u2',main = 'gumbel - theta=50')
```

Clayton Copula

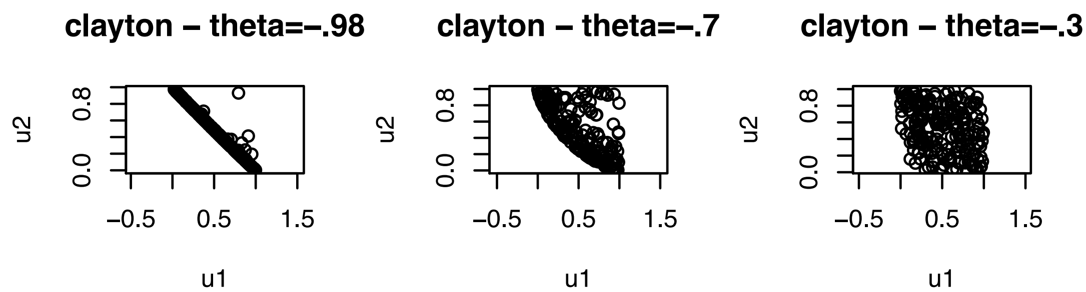
$\phi(u) = u^{-\theta} - 1$

- General d-dimensional Clayton copula cdf:

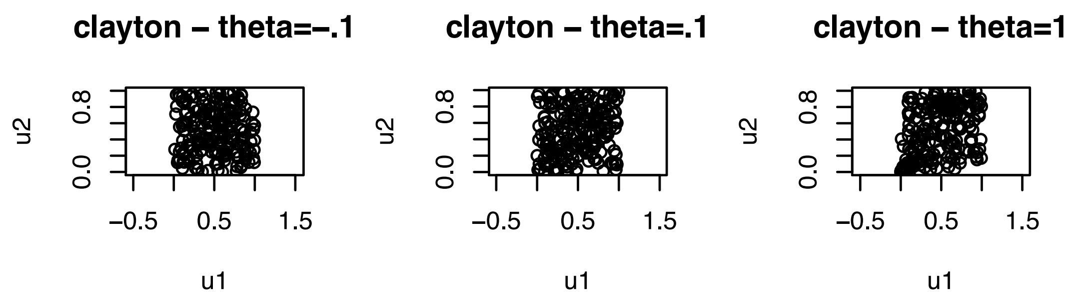
$$C(u_1, \ldots, u_d) = (u_1^{-\theta} + \cdots + u_d^{-\theta} - d + 1)^{\frac{-1}{\theta}}, \quad \theta > 0$$

- Independence copula when $\underline{\theta \to 0}$
- Define clayton copula as independence copula when $\theta = 0$

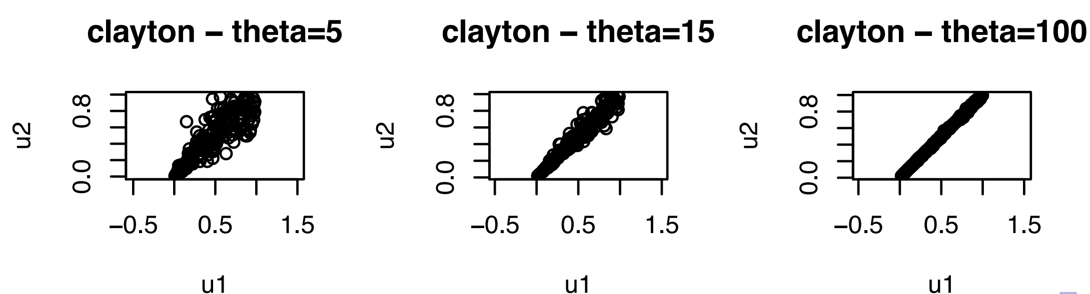
Clayton Scatter Plots



theta趋近于0,dependency逐渐消失



同样是asymmetric model



```
library(copula)
par(mfrow=c(3,3))
c <- claytonCopula(-.98, dim = 2)
uv <- rCopula(100, c)
plot(uv[,1],uv[,2],xlab='u1',ylab='u2',asp=1,main = 'clayton - theta=-.98')
c <- claytonCopula(-.7, dim = 2)
uv <- rCopula(100, c)
plot(uv[,1],uv[,2],xlab='u1',ylab='u2',asp=1,main = 'clayton - theta=-.7')
c <- claytonCopula(-.3, dim = 2)
uv <- rCopula(100, c)
plot(uv[,1],uv[,2],xlab='u1',ylab='u2',asp=1,main = 'clayton - theta=-.3')
c <- claytonCopula(-.1, dim = 2)
uv <- rCopula(100, c)
plot(uv[,1],uv[,2],xlab='u1',ylab='u2',asp=1,main = 'clayton - theta=-.1')
c <- claytonCopula(.1, dim = 2)
uv <- rCopula(100, c)
plot(uv[,1],uv[,2],xlab='u1',ylab='u2',asp=1,main = 'clayton - theta=.1')
c <- claytonCopula(1, dim = 2)
uv <- rCopula(100, c)
plot(uv[,1],uv[,2],xlab='u1',ylab='u2',asp=1,main = 'clayton - theta=1')
```

```
c <- claytonCopula(5, dim = 2)
uv <- rCopula(100, c)
plot(uv[,1],uv[,2],xlab='u1',ylab='u2',asp=1,main = 'clayton - theta=5')
c <- claytonCopula(15, dim = 2)
uv <- rCopula(100, c)
plot(uv[,1],uv[,2],xlab='u1',ylab='u2',asp=1,main = 'clayton - theta=15')
c <- claytonCopula(100, dim = 2)
uv <- rCopula(100, c)
plot(uv[,1],uv[,2],xlab='u1',ylab='u2',asp=1,main = 'clayton - theta=100')
```

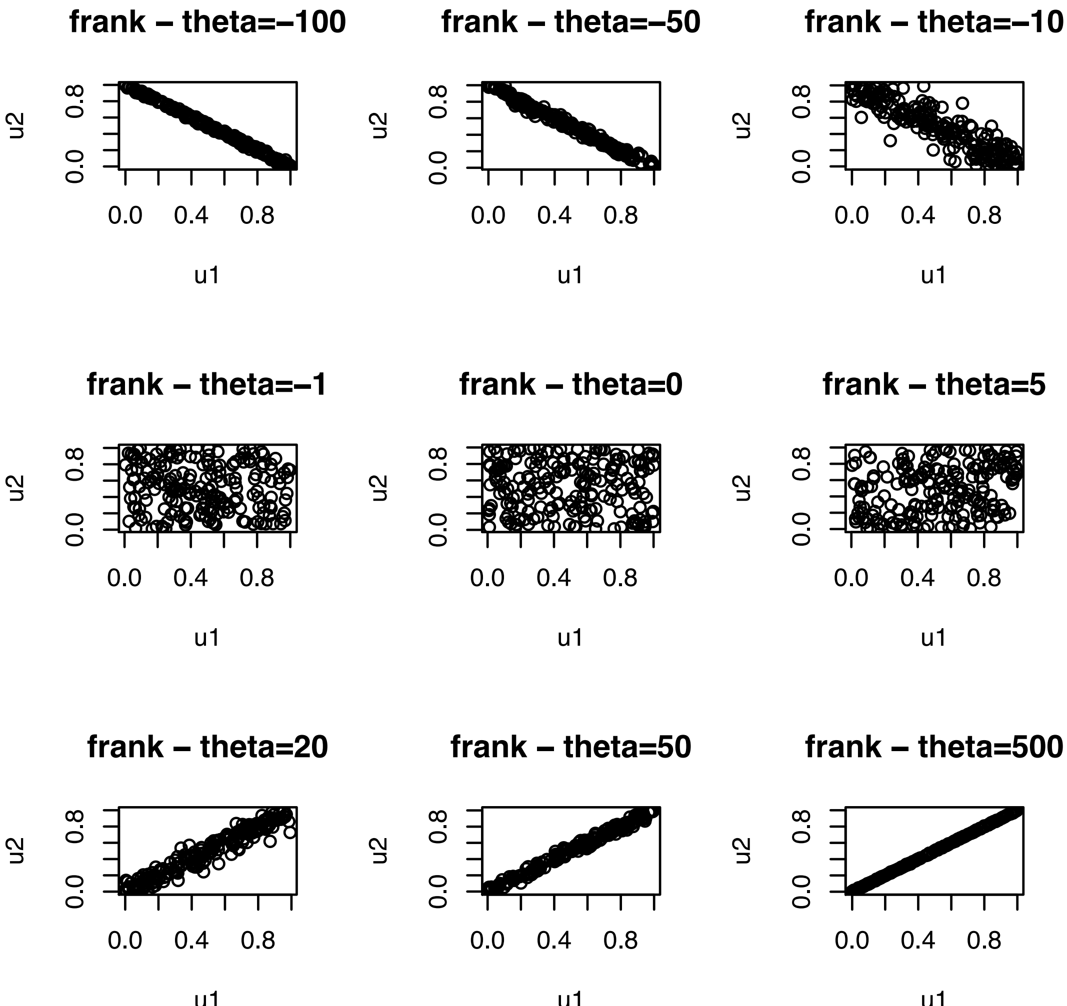
Frank Copula

- Bivariate cdf for Frank copula model:

$$C(u_1, u_2) = -\frac{1}{\theta} \log \left[1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1} \right] \qquad -\infty < \theta < \infty$$

- Independence copula when $\theta \to 0$ - define Frank as the independent copula when $\theta = 0$ 需要专门定义 $\theta = 0$

Frank Scatter Plots



```
> library(copula)
> par(mfrow=c(3,3))
> c <- frankCopula(-100, dim = 2)
> uv <- rCopula(200, c)
> plot(uv[,1],uv[,2],xlab=’u1’,ylab=’u2’,main = ’frank - theta=-100’)
> c <- frankCopula(-50, dim = 2)
> uv <- rCopula(200, c)
> plot(uv[,1],uv[,2],xlab=’u1’,ylab=’u2’,main = ’frank - theta=-50’)
> c <- frankCopula(-10, dim = 2)
> uv <- rCopula(200, c)
> plot(uv[,1],uv[,2],xlab=’u1’,ylab=’u2’,main = ’frank - theta=-10’)
> c <- frankCopula(-1, dim = 2)
> uv <- rCopula(200, c)
> plot(uv[,1],uv[,2],xlab=’u1’,ylab=’u2’,main = ’frank - theta=-1’)
> c <- frankCopula(0, dim = 2)
> uv <- rCopula(200, c)
> plot(uv[,1],uv[,2],xlab=’u1’,ylab=’u2’,main = ’frank - theta=0’)
> c <- frankCopula(1, dim = 2)
> uv <- rCopula(200, c)
> plot(uv[,1],uv[,2],xlab=’u1’,ylab=’u2’,main = ’frank - theta=5’)
> c <- frankCopula(20, dim = 2)
> uv <- rCopula(200, c)
> plot(uv[,1],uv[,2],xlab=’u1’,ylab=’u2’,main = ’frank - theta=20’)
> c <- frankCopula(50, dim = 2)
```

Multivariate Estimation Using Copula/Utility

- Want "end-to-end" joint estimation approach
  - Maximum-likelihood (joint) vs. Pseudo-likelihood
- Assessing goodness of model

  for copula
  - Comparison of bivariate empirical cdf's with estimated bivariate cdf's (with potential focus on tails)

    graphical
  - Model selection techniques (AIC and BIC)
- Utility: Assessing portfolios $\sum_{i=1}^{p} w_i R_i$ - utilize

用蒙特卡洛模拟计算VaR和ES来验证

Model Selection: AIC and BIC

- In many problems, would like quantitative general theory for model selection
  - Getting more complicated models with more parameters will always potentially give a better fit to the data – question is how to tell
  - Examples:
    - Linear regression and choosing the independent variables
    - Multivariate modeling and choosing between models (on marginals and dependency)
- AIC criteria: choose model which minimizes

$$\text{AIC} = -2\ell\left(\hat{\boldsymbol{\theta}}_{ML}\right) + 2p$$

可以用来比较t和normal分布

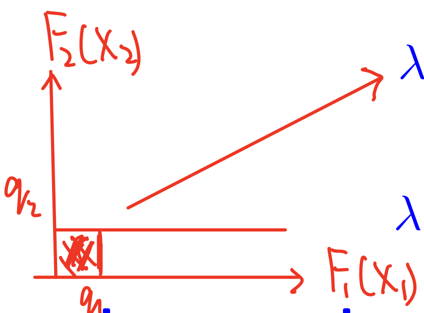
- BIC Criteria: choose model which minimizes

$$\text{BIC} = -2\ell\left(\hat{\boldsymbol{\theta}}_{ML}\right) + \boxed{\log(n)p}$$

  - ℓ is log likelihood
  - p is number of parameters
  - $\hat{\boldsymbol{\theta}}_{ML}$ is MLE estimate
  - n is sample size

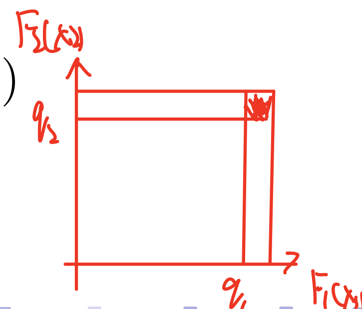
Tail Dependence (Section 8.7)

- For multivariate models, there is interest in tail dependency
  - Example: With X_1, X_2 being relative returns of two assets, want to know probability of having large loss simultaneously in both
  - Answer can be quite different between the various models and has implications for assessing portfolio VaR
- Definition: With X_1, X_2 above, the **lower and upper coefficients of tail dependence** are defined as



$$\lambda_\ell = \lim_{q \downarrow 0} P(F_1(X_1) \le q \mid F_2(X_2) \le q)$$

$$\lambda_{\mathbf{u}} = \lim_{q \uparrow 1} P(F_1(X_1) \ge q \mid F_2(X_2) \ge q)$$



Interpretation

Tail Dependency

Example. Suppose X_1, X_2 are as on previous slide with copula model C.

(a) Derive expressions for the lower and upper coefficients of tail dependency in terms of the copula model.

(b) Assuming X_1, X_2 are independent, derive the lower and upper coefficients of tail dependency.

(c) Assuming X_1, X_2 have perfect co-monotonicity, derive the lower and upper coefficients of tail dependency.

Answer. (a) $\lambda_l = \lim_{q \to 0} P[F_1(X_1) \le q \mid F_2(X_2) \le q] = \lim_{q \to 0} \frac{P[F_1(X_1) \le q \cap F_2(X_2) \le q]}{P[F_2(X_2) \le q]}$

$= \lim_{q \to 0} \frac{C(q,q)}{q}$

$\lambda_u = \lim_{q \to 1} 2 - \frac{1 - C(q,q)}{1-q}$

(b) $C(q,q) = q^2 \longrightarrow \begin{cases} \lambda_l = 0 \\ \lambda_u = 0 \end{cases}$

(c) |

Copula Tail Dependency - Gaussian and t

- For Gaussian copula, turns out that

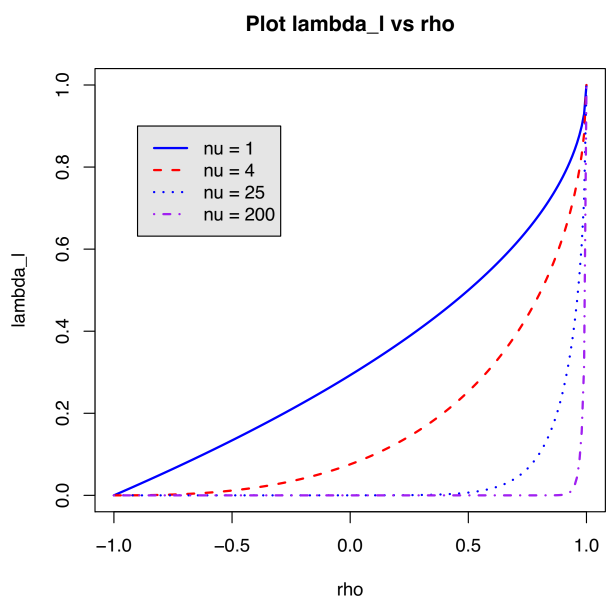
$$\underline{\lambda_\ell = \lambda_u = 0}$$

Use t distribution to catch tail dependence

- For t-copula with df ν and correlation ρ,

$$\boxed{\lambda_u = \lambda_\ell = 2F_{t,\nu+1}\left[-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right]}$$

where $F_{t,\nu+1}$ is the cdf of the t distn with df, $\nu + 1$



- Note behavior when $\nu \to \infty$
  - justified since ....

R-code for Plot

```
> # Tail dependency plot
> rhov = .99999999*seq(-1,1,.01)
> nuv = c(1,4,25,200)
> n = length(rhov)
> nnu = length(nuv)
> lambdau = matrix(0,nnu,n)
> for (j in 1:nnu){
+ for (i in 1:n){
+ lambdau[j,i] = 2*pt(-sqrt((nuv[j]+1)*(1-rhov[i])/(1+rhov[i])),
+ nuv[j]+1)
+ }
+ }
> windows()
> plot(rhov,lambdau[1,],xlim=c(-1,1),ylim=c(0,1),xlab='rho',ylab=
+ 'lambda_l',type='l',lwd=2,col='blue',main='Plot lambda_l vs rho')
> lines(rhov,lambdau[2,],lty=2,lwd=2,col='red')
> lines(rhov,lambdau[3,],lty=3,lwd=2,col='blue')
> lines(rhov,lambdau[4,],lty=4,lwd=2,col='purple')
> legend(-.9,.9, c("nu = 1","nu = 4","nu = 25","nu = 200"),lty=c(1,2,
+ 3,4),lwd=c(2,2,2,2),col=c("blue","red","blue","purple"), bg="gray90")
```

Example: Joint Copula Estimation

Data: IBM and CRSP daily return data in Ecdat library

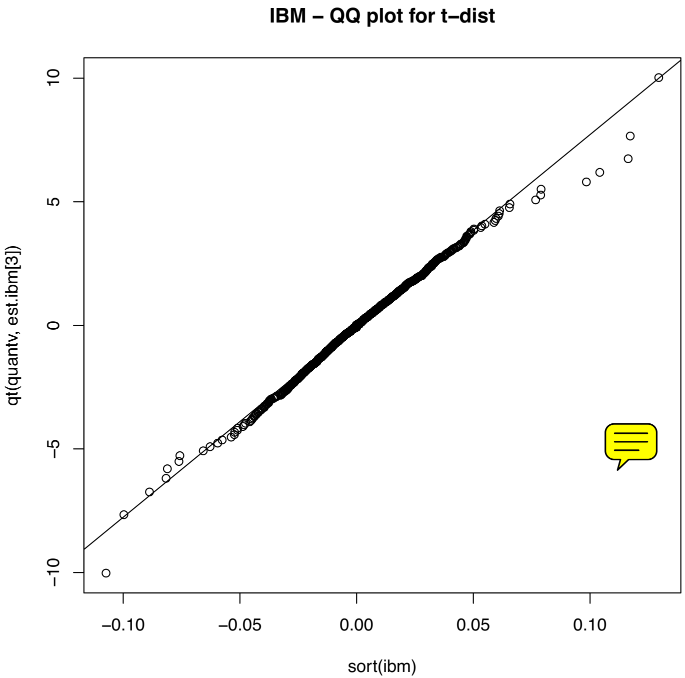
Objective 1: Fit a joint distribution to the data via copula method

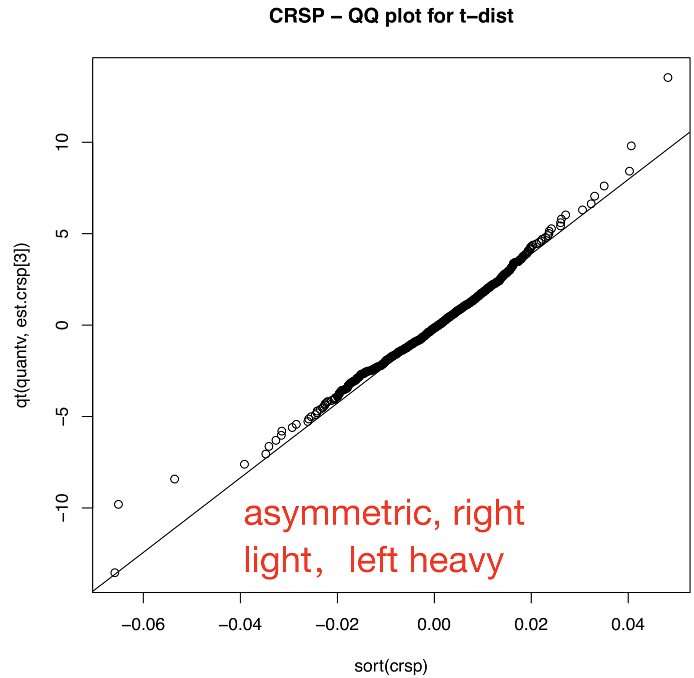
- Recall the 3 steps

Objective 2: Utilize the joint model to find minimum VaR at $\alpha = .005$ Search w

```
> library(Ecdat) # need for the data
> library(copula) # for copula functions
> library(fGarch) # need for standardized t density
> library(MASS) # need for fitdistr and kde2d
> library(fCopulae) # additional copula functions (pempiricalCopula
>                   # and ellipticalCopulaFit)
> data(CRSPday,package="Ecdat")
> ibm = CRSPday[,5]
> crsp = CRSPday[,7]
> est.ibm = as.numeric(fitdistr(ibm,"t")$estimate)
> est.crsp = as.numeric(fitdistr(crsp,"t")$estimate)
> est.ibm
[1] 0.0002936796 0.0126959301 4.2761559095
> est.crsp
[1] 0.0009034741 0.0052196399 3.4739822923
> N = length(ibm)
> quantv = (1/N)*seq(.5,N-.5,1)
> qqplot(sort(ibm),qt(quantv,est.ibm[3]),
  + main='IBM - QQ plot for t-dist')
> abline(lm(qt(c(.25,.75),est.ibm[3])~quantile(ibm,c(.25,.75))))
> qqplot(sort(crsp),qt(quantv,est.crsp[3]),
  + main='CRSP - QQ plot for t-dist')
> abline(lm(qt(c(.25,.75),est.crsp[3])~quantile(crsp,c(.25,.75))))
```

QQ Plots of Marginal distributional estimates





Estimation of Copulas - Example R-code

```
> # Need to convert to standard deviation for incorporating within "ptsd"
> est.ibm[2] = est.ibm[2]*sqrt(est.ibm[3]/(est.ibm[3]-2))
> est.crsp[2] = est.crsp[2]*sqrt(est.crsp[3]/(est.crsp[3]-2))
>
> # Generating initial estimate of correlation for t-copula
> cor_tau = cor(ibm,crsp,method="spearman")
> omega = cor_tau
> omega
[1] 0.4735411
> n = length(ibm)
> data1 = cbind(pstd(ibm,mean=est.ibm[1],sd=est.ibm[2],nu=est.ibm[3]),
+ pstd(crsp,mean=est.crsp[1],sd=est.crsp[2],nu=est.crsp[3]))
>
> cop_t_dim2 = tCopula(omega, dim = 2, dispstr = "un", df = 4)
> ft1 = fitCopula(cop_t_dim2, optim.method="L-BFGS-B", data=data1,
+ start=c(omega,5),lower=c(0,2.5),upper=c(.5,15) )
> ft1
The estimation is based on the maximum pseudo-likelihood
and a sample of size 2528.
       Estimate Std. Error  z value Pr(>|z|)
rho.1 0.4936568 0.02459639 20.07029        0
df    9.8536751         NA       NA       NA
The maximized loglikelihood is  361.9846
The convergence code is  0
```

Estimation of Copulas - Example R-code

```
> fnorm = fitCopula(data=data1,copula=normalCopula(.3,dim=2),
+ optim.method="BFGS",start=.5)
> fnorm
The estimation is based on the maximum pseudo-likelihood
and a sample of size 2528.
      Estimate Std. Error  z value Pr(>|z|)
rho.1 0.490532 0.01352036 36.28098        0
The maximized loglikelihood is  347.1844
The convergence code is  0
> fgumbel = fitCopula(data=data1,optim.method="BFGS",
+ copula=gumbelCopula(3,dim=2),start=2)
> fgumbel
The estimation is based on the maximum pseudo-likelihood
and a sample of size 2528.
      Estimate Std. Error  z value Pr(>|z|)
param 1.430171 0.02222564 64.34779        0
The maximized loglikelihood is  313.2257
The convergence code is  0
```

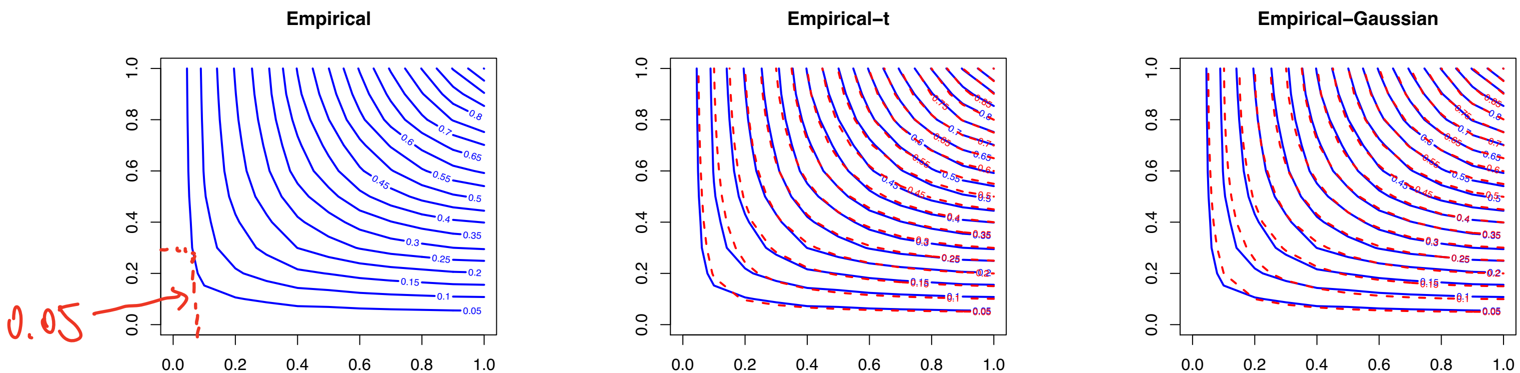
Estimation of Copulas - Example R-code

```
> ffrank = fitCopula(data=data1,optim.method="BFGS",
+ copula=frankCopula(3,dim=2),start=1)
> ffrank
The estimation is based on the maximum pseudo-likelihood
and a sample of size 2528.
      Estimate Std. Error  z value Pr(>|z|)
param 3.301376   0.136598 24.16856        0
The maximized loglikelihood is  325.2867
The convergence code is  0
> fclayton = fitCopula(data=data1,optim.method="BFGS",
  + copula=claytonCopula(1,dim=2),start=1)
> fclayton
The estimation is based on the maximum pseudo-likelihood
and a sample of size 2528.
      Estimate Std. Error  z value Pr(>|z|)
param 0.704918 0.03292131 21.41221        0
The maximized loglikelihood is  293.1102
The convergence code is  0
```

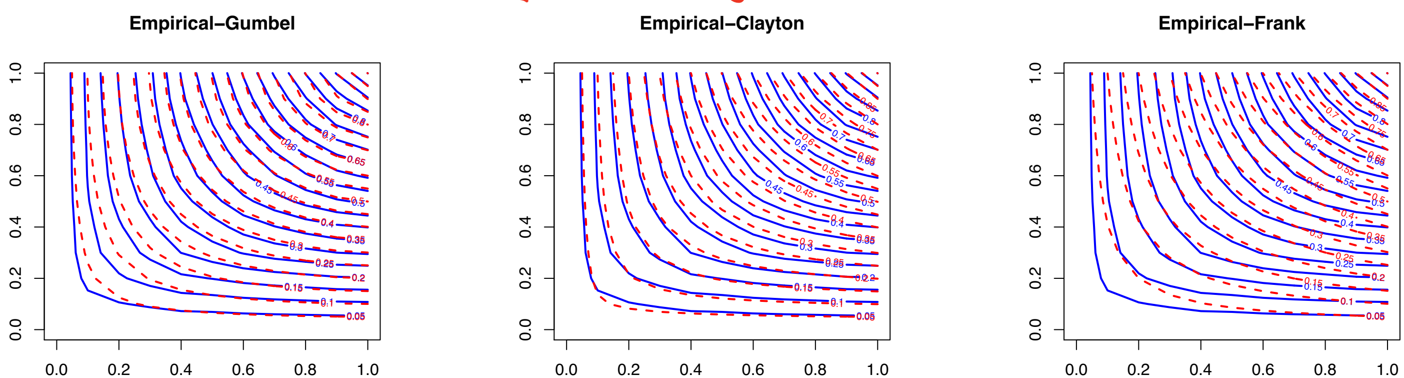
Diagnostic Plots for Copulas Estimation

bivariant CDF 等高线 plot

- Solid-blue - empirical cdf & dashed-red - parametric copula cdf



$$\hat{F}_n(u,v) = \#\{(F_X(x), F_Y(y)) \le (u,v)\}$$



Diagnostic Plots for Copulas Estimation - Example R-code

```
> u1 = data1[,1]
> u2 = data1[,2]
> dem = pempiricalCopula(u1,u2)
>
> contour(dem$x,dem$y,dem$z,main="Empirical",col='blue',lty=1,lwd=2,nlevel=20)

> contour(dem$x,dem$y,dem$z,main="Empirical-t",col='blue',lty=1,lwd=2,nlevel=20)
> ct = tCopula (ft1@estimate[1], dim = 2, dispstr = "un", df = ft1@estimate[2])
> utdis = rCopula(100000,ct)
> demt = pempiricalCopula(utdis[,1],utdis[,2])
> contour(demt$x,demt$y,demt$z,main="t",col='red',lty=2,lwd=2,add=TRUE,
+ nlevel=20)

> contour(dem$x,dem$y,dem$z,main="Empirical-Gaussian",col='blue',lty=1,lwd=2,
+ nlevel=20)
> cn <- normalCopula(fnorm@estimate[1], dim = 2, dispstr = "un")
> utdis = rCopula(100000,cn)
> demt = pempiricalCopula(utdis[,1],utdis[,2])
> contour(demt$x,demt$y,demt$z,main="Gauss",col='red',lty=2,lwd=2,add=TRUE,
+ nlevel=20)
```

```
> contour(dem$x,dem$y,dem$z,main="Empirical-Gumbel",col='blue',lty=1,lwd=2,
+ nlevel=20)
> cg <- gumbelCopula(fgumbel@estimate, dim = 2)
> utdis = rCopula(100000,cg)
> demt = pempiricalCopula(utdis[,1],utdis[,2])
> contour(demt$x,demt$y,demt$z,col='red',lty=2,lwd=2,add=TRUE,nlevel=20)

> contour(dem$x,dem$y,dem$z,main="Empirical-Clayton",col='blue',
+ lty=1,lwd=2,nlevel=20)
> cc <- claytonCopula(fclayton@estimate, dim = 2)
> utdis = rCopula(100000,cc)
> demt = pempiricalCopula(utdis[,1],utdis[,2])
> contour(demt$x,demt$y,demt$z,col='red',lty=2,lwd=2,add=TRUE,
+ nlevel=20)

> contour(dem$x,dem$y,dem$z,main="Empirical-Frank",col='blue',
+ lty=1,lwd=2,nlevel=20)
> cf <- frankCopula(fclayton@estimate, dim = 2)
> utdis = rCopula(100000,cf)
> demt = pempiricalCopula(utdis[,1],utdis[,2])
> contour(demt$x,demt$y,demt$z,col='red',lty=2,lwd=2,add=TRUE,
+ nlevel=20)
```

Copula-based VaR

Objective: Based on results from IBM and CRSP, utilize the copula estimation for doing VaR at α

- Step 1: Simulate $(U_1, V_1), \ldots, (U_n, V_n)$ from copula model C, and then utilize the marginal quantile functions, i.e.,

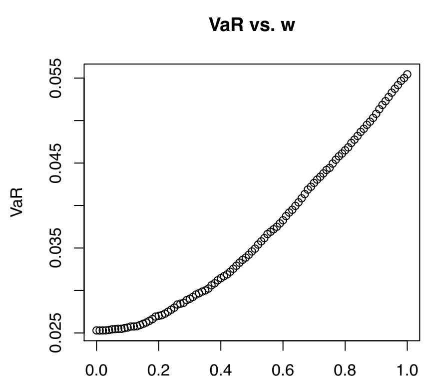
$$(X_i, Y_i) = (F_1^{-1}(U_i), F_2^{-1}(V_i))$$

  Exercise Show that the distribution of above corresponds to the copula model.

- Step 2: For weighting $0 \leq w \leq 1$, compute the α-quantile of $\{wX_i + (1-w)Y_i\}_1^n$, and optimize over w.

Example Calculation of Copula-Based VaR

```
> ct = tCopula (ft1@estimate[1], dim = 2, dispstr = "un", df = ft1@estimate[2])
> uvsim = rCopula(100000,ct)
> w = seq(0,1,.01)
> n = length(w)
> VaRv = rep(0,n)
> data_sim = cbind(qstd(uvsim[,1],mean=est.ibm[1],sd=est.ibm[2],nu=est.ibm[3]),
  + qstd(uvsim[,2],mean=est.crsp[1],sd=est.crsp[2],nu=est.crsp[3]))
>
> for (i in 1:n){
+ datat = w[i]*data_sim[,1] + (1-w[i])*data_sim[,2];
+ VaRv[i] = -quantile(datat,.005)
+ }
> plot(w,VaRv,xlab='w',ylab='VaR',main="VaR vs. w")
> wmax = w[which.min(VaRv)]
> VaR = VaRv[which.min(VaRv)]
> c(wmax,VaR)
[1] 0.00000000 0.02540281
```



由于IBM和CRSP之间有很强的相关性,拟合结果说明w=0,只投资CRSP

AIC Criteria

- Recall that the bivariate pdf model is given by

$$f(x_1, x_2; \theta_1, \theta_2, \theta_c) = c_\theta(F_{1\theta_1}(x_1), F_{2,\theta_2}(x_2))f_{1,\theta_1}(x_1)f_{2,\theta_2}(x_2)$$

- For the pseudo log-likelihood function approach (i.e., estimating the marginals first), utilizing the AIC criteria to compare copula models amounts to focusing on log-likelihood of copula estimation alone. Justify.

```
> AIC_t = -2*ft1@loglik+2*2
> AIC_gauss = -2*fnorm@loglik+2*1
> AIC_gumbel = -2*fgumbel@loglik+2*1
> AIC_clayton = -2*fclayton@loglik+2*1
> AIC_frank = -2*ffrank@loglik+2*1
> c(AIC_t,AIC_gauss,AIC_gumbel,AIC_clayton,AIC_frank)
[1] -719.9693 -692.3688 -624.4514 -584.2204 -648.5734
```

t最好