Risk Management — Assignment 4

AJ*

Autumn 2021

Question 1

```
## import data
prices <- read.table(file="RMIP Data Tutorial.txt", header=TRUE)
prices$Date <- as.Date(prices$Date, '%m/%d/%Y')

## log returns
lr.NL <- diff(log(prices$Netherlands)); lr.GR <- diff(log(prices$Greece))
lr <- cbind(lr.NL, lr.GR)
```

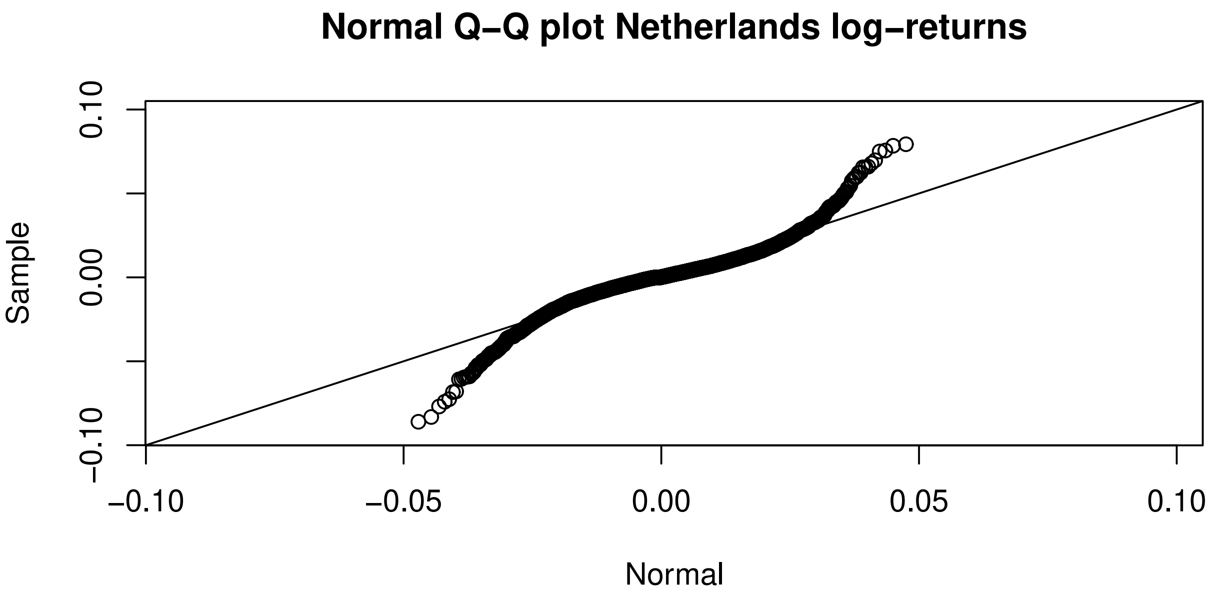
Question 2

```
## a) Jarque-Bera test statistic NL
n <- length(lr.NL) ## 5217
x <- lr.NL - mean(lr.NL)
m2 <- mean(x^2); m3 <- mean(x^3); m4 <- mean(x^4)
b <- (m3^2) / (m2^3)
k <- m4 / (m2^2)
jb <- 1/6 * n * (b + 1/4*(k-3)^2)
> 1-pchisq(jb, df=2) ## p-value
[1] 0

## b) Jarque-Bera test statistic GR
n <- length(lr.GR)
x <- lr.GR - mean(lr.GR)
m2 <- mean(x^2); m3 <- mean(x^3); m4 <- mean(x^4)
b <- (m3^2) / (m2^3)
k <- m4 / (m2^2)
jb <- 1/6 * n * (b + 1/4*(k-3)^2)
> 1-pchisq(jb, df=2) ## p-value
[1] 0

## c) We have strong grounds to reject the null hypothesis that the data
      is normally distributed (for both countries).
```

*Student number: ∞

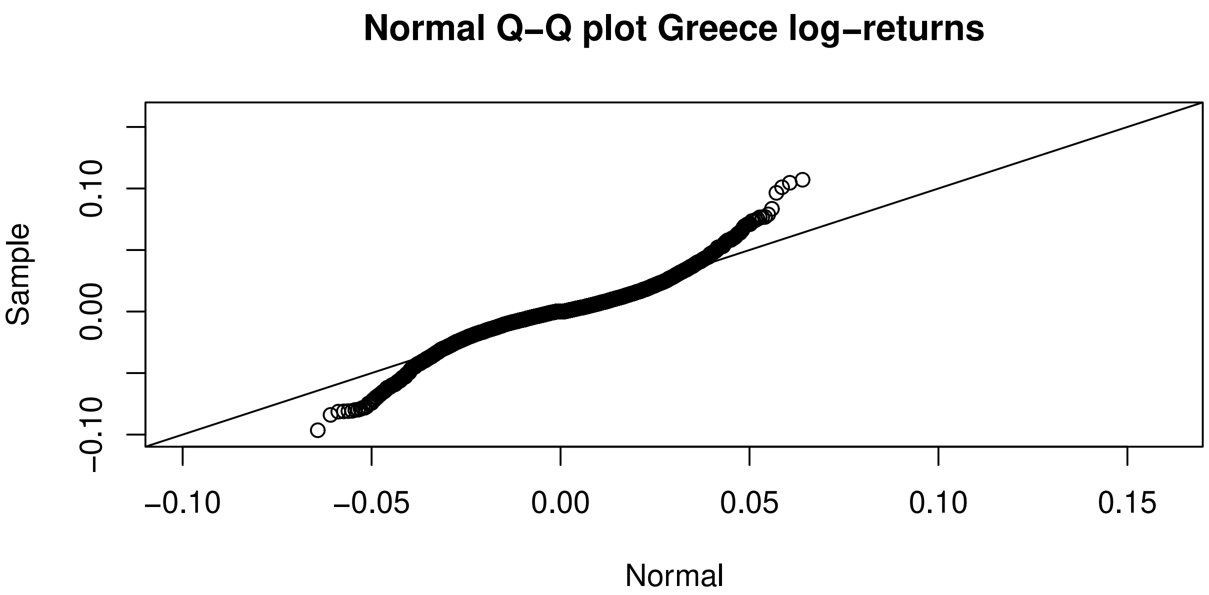
Question 3

```
## a)
## normal Q-Q plot for the NL log-returns
samp.quant.NL <- quantile(lr.NL, probs = seq(0, 1, 1/n))
norm.quant.NL <- qnorm(seq(0, 1, 1/n), mean=mean(lr.NL), sd=sd(lr.NL))
plot(norm.quant.NL, samp.quant.NL, xlim=range(samp.quant.NL),
     main="Normal Q-Q plot Netherlands log-returns", xlab = "Normal",
     ylab = "Sample")
abline(0,1)
```



The plot indicates that the sample quantiles differ significantly from the theoretical quantiles under the assumption of normality, particularly in the tails of the distribution. We therefore reject the assumption of normality.

```
## 3b)
## normal Q-Q plot for the GR log-returns
samp.quant.GR <- quantile(lr.GR, probs = seq(0, 1, 1/n))
norm.quant.GR <- qnorm(seq(0, 1, 1/n), mean=mean(lr.GR), sd=sd(lr.GR))
plot(norm.quant.GR, samp.quant.GR, xlim=range(samp.quant.GR),
     main="Normal Q-Q plot Greece log-returns", xlab = "Normal",
     ylab = "Sample")
abline(0,1)
```



Based on the same reasoning as in *(3a)*, we conclude that the observations are not normally distributed.

Question 4

We run:

```
> round(MardiaTest(lr), 5)
       K3 K3 p-value         K4 K4 p-value
  0.06625    0.00000   19.90611    0.00000
```

The statistics computed are the d-dimensional analogues of skewness and kurtosis respectively.

The p-values associated with these statistics indicate that the assumption of multivariate normality should be rejected.

Question 5

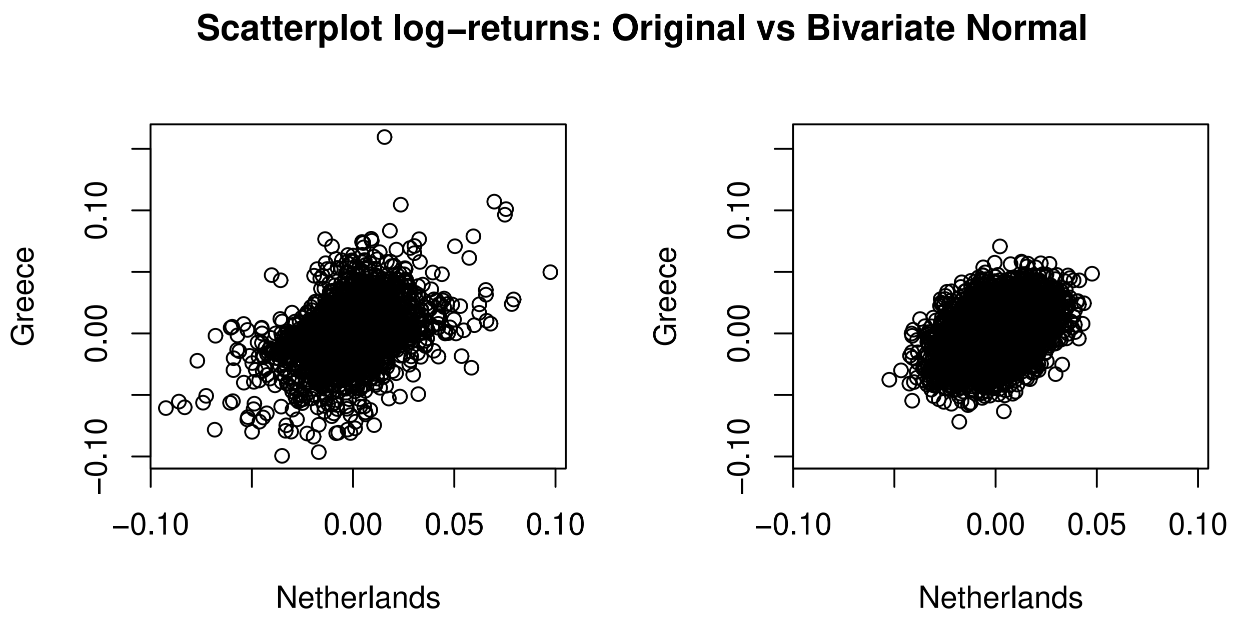
We have:

```
> fit.norm(lr)
$mu
        lr.NL         lr.GR
 0.0001686451 -0.0001013349

$Sigma
             lr.NL        lr.GR
lr.NL 1.773521e-04 8.993128e-05
lr.GR 8.993128e-05 3.261535e-04

$cor
          lr.NL     lr.GR
lr.NL 1.0000000 0.3739225
lr.GR 0.3739225 1.0000000

$ll.max
[1] 29060.16

## extract mean and covariance matrix

bivariate.lr <- fit.norm(lr)
mu <- as.numeric(bivariate.lr$mu)
Sigma <- as.matrix(bivariate.lr$Sigma)

## simulated bivariate normal log-returns

norm.lr <- rmnorm(n, mu, Sigma)
colnames(norm.lr) <- c("norm.lr.NL", "norm.lr.GR")
```

Question 6

Running the code provided in the instructions with a few modifications, we have:

```
par(mfrow=c(1,2))
plot(lr, xlab="Netherlands", ylab="Greece")
plot(norm.lr, xlim=range(lr.NL), ylim=range(lr.GR), xlab="Netherlands",
     ylab="Greece")
mtext("Scatterplot log-returns: Original vs Bivariate Normal",
      line=-2, font=2, cex=1.2, outer=TRUE)
```



We notice that both data sets display some moderate positive correlation (by design in the second case). However, the original data is far more dispersed (particularly towards the tails of the distribution) than would be expected under bivariate normality, where we have a well-defined elliptical region.

Question 7

a) `lr.ranks.NL <- rank(lr.NL)`

b) `lr.ranks.GR <- rank(lr.GR)`

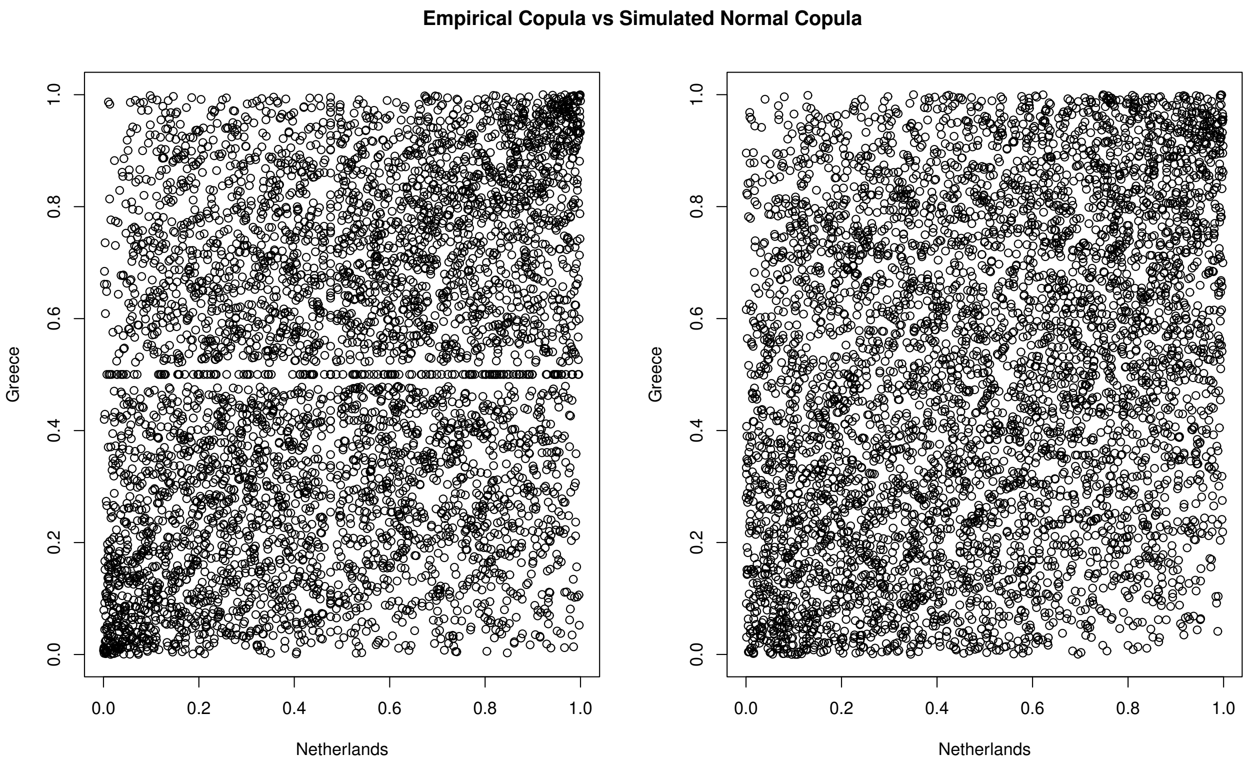
Question 8

```
> cor(lr.ranks)
            lr.ranks.NL lr.ranks.GR
lr.ranks.NL   1.0000000   0.3104008
lr.ranks.GR   0.3104008   1.0000000

> Spearman(lr)
          lr.NL     lr.GR
lr.NL 1.0000000 0.3104008
lr.GR 0.3104008 1.0000000
```

Both of these lines produce the same output, which is unsurprising given we have computed the correlation of the ranked observations (scaled to the interval $[0, 1]$) and compared them to Spearman's rank correlation, which is by definition the same quantity.

Question 9

```
## a)
par(mfrow=c(1,2))
plot(lr.ranks, xlab="Netherlands", ylab="Greece")
plot(norm.lr.ranks,  xlab="Netherlands", ylab="Greece")
mtext("Empirical Copula vs Simulated Normal Copula",
      line=-2, font=2, cex=1.2, outer=TRUE)
```

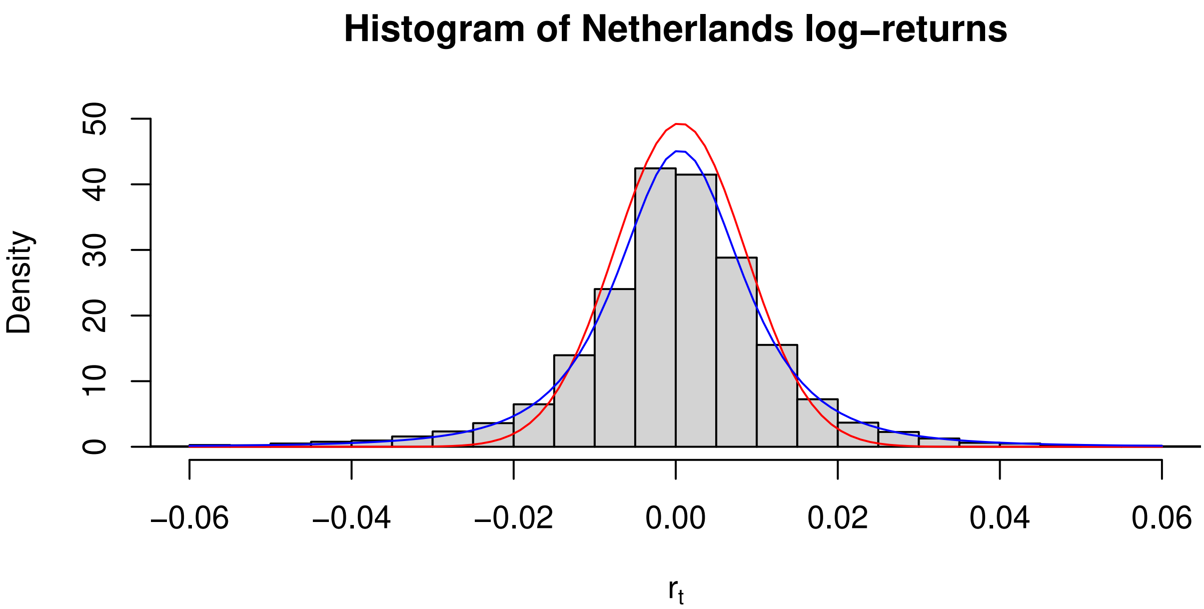


The normal copula appears to be acceptable, however, we note that the empirical copula displays some clustering in the lower-left and upper-right corners.

(b) The observations under consideration are the closing prices of a market index, and markets do not trade on national holidays (in most cases). Hence, the closing prices on consecutive days are necessarily the same, and the log-return is zero. Assuming that the distribution is symmetric around zero, we would expect these points to be near the 50$^{\text{th}}$ percentile of the distribution *i.e.* the median.

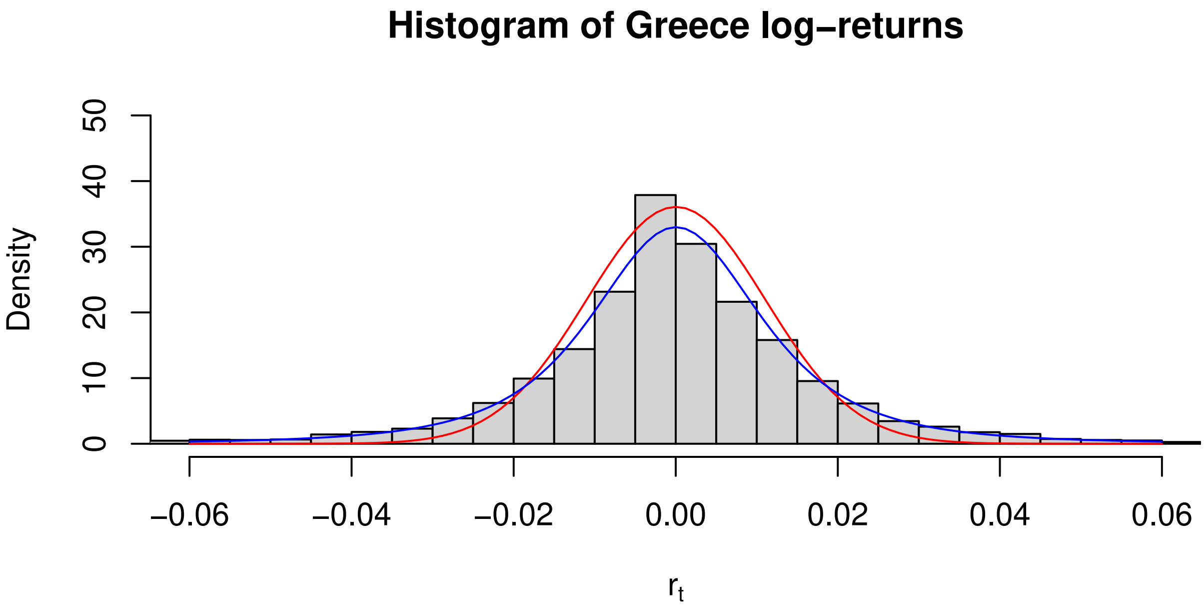
Question 10

```
## a)
t.NL <- fit.st(lr.NL)
nu.NL <- t.NL$par.ests[1]
mu.NL <- t.NL$par.ests[2]
sigma.NL <- t.NL$par.ests[3]

## b)
t.GR <- fit.st(lr.GR)
nu.GR <- t.GR$par.ests[1]
mu.GR <- t.GR$par.ests[2]
sigma.GR <- t.GR$par.ests[3]

## c)
dst <- function(x,nu,mu,sigma){1/sigma*dt((x-mu)/sigma, nu)}
hist(lr.NL, prob=T, breaks=40, xlim=c(-0.06,0.06), ylim=c(0,50),
     main="Histogram of Netherlands log-returns", xlab = expression(r[t]))
curve(dnorm(x, mean=mu.NL, sd=sigma.NL), add=T, col="red")
curve(dst(x, nu.NL, mu.NL, sigma.NL), add=T, col="blue")
```



The Student t density appears to be the better fit, given that it allocates more density to the tails of the distribution.

```
## d)
hist(lr.GR, prob=T, breaks=40, xlim=c(-0.06,0.06), ylim=c(0,50),
     main="Histogram of Greece log-returns", xlab = expression(r[t]))
curve(dnorm(x, mean=mu.GR, sd=sigma.GR), add=T, col="red")
curve(dst(x, nu.GR, mu.GR, sigma.GR), add=T, col="blue")
```



As in *(10c)*, the Student t density appears to be the better fit.

Question 11

Consider $X = F_1^{-1}(U)$, $Y = F_2^{-1}(V)$ and $F_{X,Y}^{-1}(X, Y)$, where U , $V \sim \text{Uniform}(0, 1)$.

We have:

$$\begin{aligned} F_X^{-1}(x) &= F_{X,Y}^{-1}(X, \infty) \\ &= \mathrm{P}[X \leq x, Y \leq \infty] \\ &= \mathrm{P}[F_1^{-1}(U) \leq x] \\ &= \mathrm{P}[U \leq F_1(x)] \\ &= F_1(x) \end{aligned}$$

where the last line follows from the the distribution function of U.

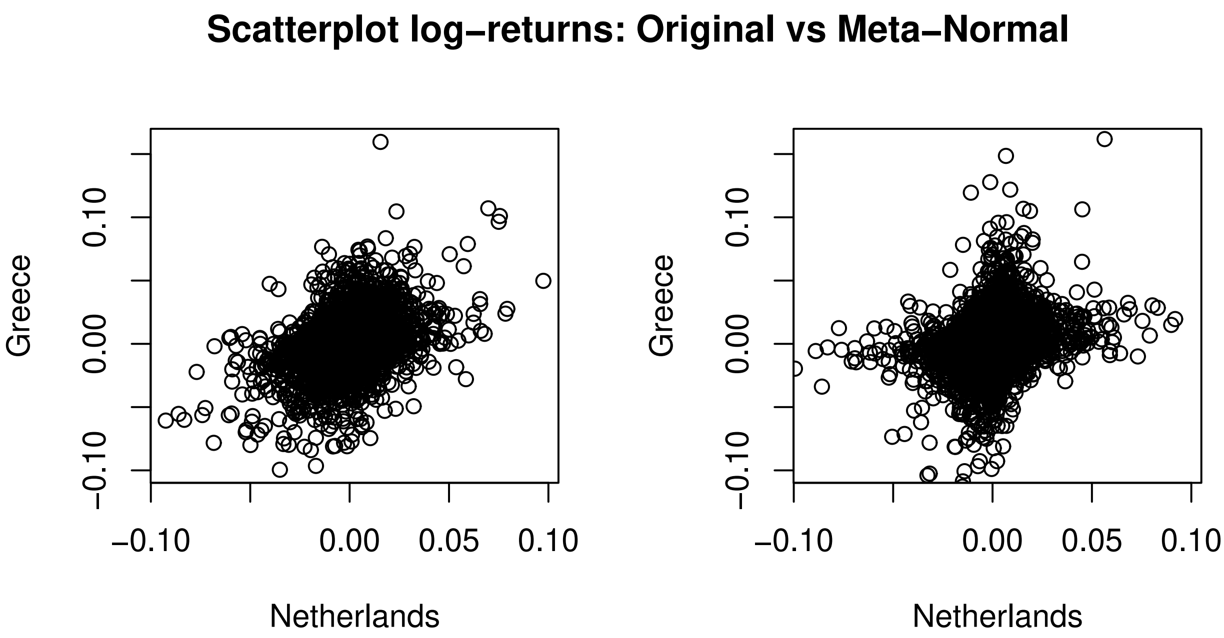
Now,

$$\begin{aligned} F(x, y) &= \mathrm{C}\,(F_1(x), F_2(y)) \quad [\textit{Sklar's Theorem Part I}] \\ &= \mathrm{C}\,(F_1(F_1^{-1}(u)), F_2(F_2^{-1}(v))) \\ &= \mathrm{C}\,(u, v). \end{aligned}$$

Question 12

We have:

```
par(mfrow=c(1,2))
plot(lr, xlab="Netherlands", ylab="Greece")
plot(meta.lr, xlim=range(lr.NL), ylim=range(lr.GR), xlab="Netherlands",
     ylab="Greece")
mtext("Scatterplot log-returns: Original vs Meta-Normal",
      line=-2, font=2, cex=1.2, outer=TRUE)
```

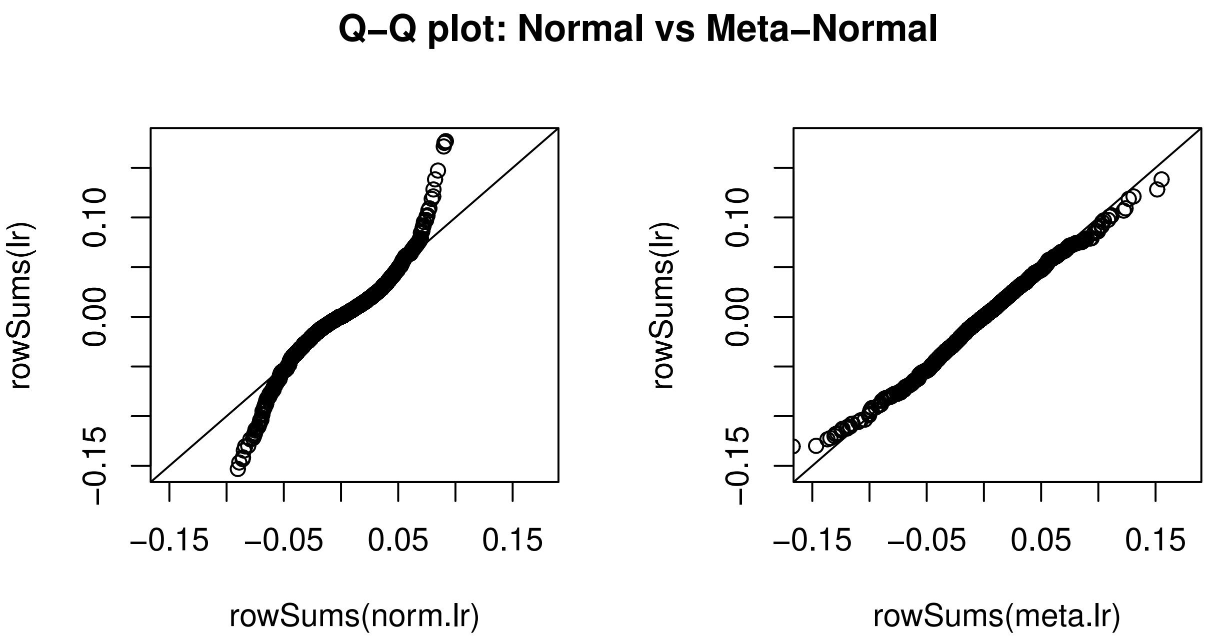


The meta-normal simulation displays a comparable level of dispersion to the original data, and contains extreme values like the original data set. This is definitely an improvement on the well-defined elliptical region under bivariate normality.

Question 13

Finally,

```
par(mfrow=c(1,2))
qqplot(rowSums(norm.lr), rowSums(lr), xlim=range(rowSums(lr))); abline(0,1)
qqplot(rowSums(meta.lr), rowSums(lr), xlim=range(rowSums(lr))); abline(0,1)
mtext("Q-Q plot: Normal vs Meta-Normal",
      line=-2, font=2, cex=1.2, outer=TRUE)
```



The sample and theoretical quantiles are (more) comparable under the meta-normal model, and we conclude that this is model provides a better fit for the aggregate risk.