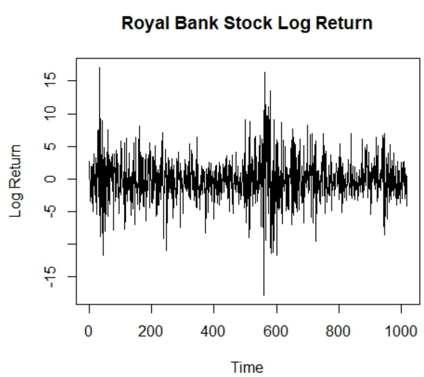
Risk Management Project - Copulas

Alice Li

July 16, 2017

1. We plot the two paths of log returns as time series.



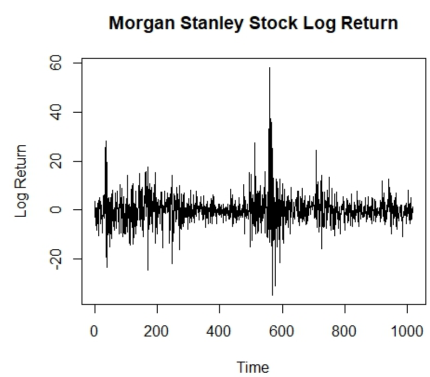


Figure 1: Royal Bank of Canada ME plot

2. Test whether Stock A and Stock B are normally distributed, using the Jarque-Bera test.

H_0: log returns of stock prices is normally distributed.
Using Jarque-Bera test from tseries package we have following results:
Royal Bank of Canada: X-squared = 272.82, df = 2, p-value = 2.2e-16 $<$ 0.01
Morgan Stanley: X-squared = 5548.8, df = 2, p-value = 2.2e-16 $<$ 0.01
Hence we reject the null hypothesis for both stocks. So they are not normally distributed.

3. Produce a scatterplot plot of the data pairs.

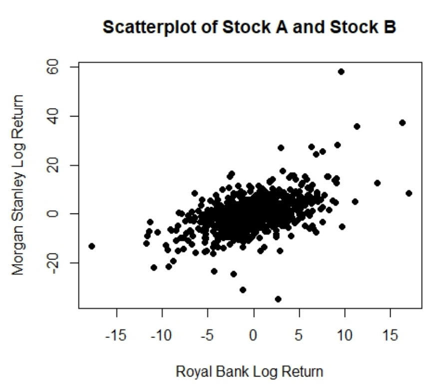


Figure 2: Q3: Scatterplot of data pairs

4. Produce a scatterplot of the probability-transformed pairs

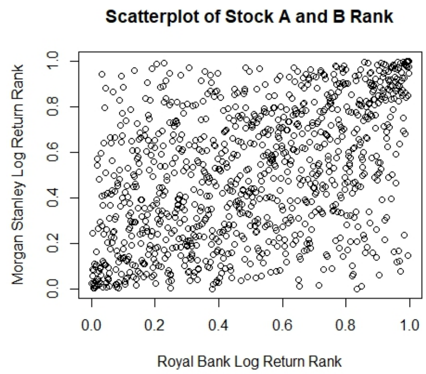


Figure 3: Q4: Scatterplot of data pairs rank

5. Study the index of upper and lower dependence.
(a) First we obtain estimates of Kendall's tau for the entire data set and also for the subsets in the upper-right and lower-left corners as the square containing the subsets gets smaller.

The Kendall's Tau for the entire data set is 0.330115.
The Kendall's Tau values for different subsets are summarized in the table below.
Let U_α represent the Kendall's Tau value for the data subset where the normalized log returns of both stocks are greater than the α-quantile returns respectively (upper-right corner). And let L_α represent the Tau value for the data subset where the normalized log returns of both stocks are less than their $(1-\alpha)$-quantile returns (lower-left corner).

α	0.70	0.80	0.95	0.99	0.995
U_α	0.2422764	0.2359596	0.3235294	0.3333333	1
$L_{1-\alpha}$	0.2393999	0.2671172	0.1029412	1	NA

Table 1: Tau Table for different data subsets

(b) Now we obtain estimates of $P(U^A > u|U^B > u)$ for $u = 0.95, 0.99, 0.995$.

u	0.95	0.99	0.995
$P(U^A > u\|U^B > u)$	0.34	0.4	0.4
$P(U^B > u\|U^A > u)$	0.34	0.4	0.4
$P(U^A < 1-u\|U^B < 1-u)$	0.34	0.2	0
$P(U^B < 1-u\|U^A < 1-u)$	0.34	0.2	0

Table 2: Conditional probabilities

(c) Comments:
The scatter plot of Rank A and B shows a positive correlation , because the data points are concentrated in the upper right and lower left corners. The positive correlation is also reflected in the positive Kendall's Tau value for entire data: 0.330115.

In part (a), we can see that in the upper-right corner (ie first row of Table 1), the Kendall's Tau estimates increase as the square containing subsets gets smaller. This implies concordance and strong upper tail dependency between two stocks. This aligns with our results in (b), as the positive non-decreasing conditional probabilities in the first 2 rows in Table 2 indicate upper tail dependence. For example, when $\alpha = 0.95, P(U^A > u|U^B > u) = 0.34$; this means that the probability that Stock A returns lies in the upper 5%-quantile is 34%, given that Stock B returns lies in the upper 5%-quantile , indicating positive upper tail dependence. However, please note that these results may not be reliable, as very few data points fall into upper right corner subsets with $\alpha = 0.99, 0.995$.

In the lower left corner(ie second row of Table 1), the Kendall's Tau estimates appear to be positive. This implies Stock A and Stock B returns are concordant, but inconclusive about lower tail dependency as there is no strictly increasing or decreasing pattern from $\alpha = 0.70$ to $\alpha = 0.995$. This aligns with (b) Table 2, as the the conditional probabilities are decreasing in the bottom 2 rows, indicating no lower tail dependence.
However, please note that these results may not be reliable, as very few data points fall into lower left corner subsets with $\alpha = 0.99, 0.995$.

6. Fitting of Copulas
We first fit Gumbel copula, Gaussian copula and Student's t copula using functions in R from different packages. Then we determine the goodness of fit by using maximum likelihood and AIC(Akaike Information Criterion). Note that Gumbel copula has 1 parameter, Gaussian copula has 1 parameter and student's t copula has 2 parameters.

Distribution	Gumbel	Gaussian	Student's t
ll.max	159.5844	149.9902	1092.222
AIC	-317.1689	-297.9803	-2180.444

From the table we can see Student's t copula has maximum log likelihood. But maybe this result is influenced by the fact that student's t copula has one more parameter than other two copulas. So we used AIC formula: $AIC = -2ll.max + 2k$ where k is number of parameters. And from AIC, student's t is still the best model. As we learned in class, student's t copula has upper and lower tail dependency, and from Kendall's Tau we know that our data set also has strong upper and lower tail dependency. This fact could help to explain why student's t copula fit the most into our data.