INDIAN INSTITUTE OF TECHNOLOGY KANPUR

MTH392A : UNDER GRADUATE PROJECT II

STUDENT'S NAME : RAHUL VERMA

SUPERVISOR : ARNAB HAZRA

Project Report : Joint modeling of the upper and lower tail dependence in daily log-return of Indigo and Spicejet airlines during the COVID-19 pandemic

---

Submitted To:
Dootika Vats (convenor)
Neeraj Misra
Amit Mitra
Shalabh
Subhra Sankar Dhar
Arnab Hazra
Dept. of Math. and Stat.

Submitted By:
Rahul Verma
190671
B.S. Mathematics and Statistics

April 15, 2022

Contents

Abstract

In early 2020, the rapid spread of the COVID-19 virus prompted the governments all across the globe to take various necessary actions including banning all international and domestic flights to prevent the spread of the virus. The curbs were reduced and reimposed several times due to the dynamic pandemic situation and this led to high volatility in the stock markets around the world. In this study, we focus on analyzing the joint volatility in the stock prices of the two largest Indian private airlines Indigo and SpiceJet, that have market shares 48.2% and 15.6%, respectively. While the majority of the literature focuses on analyzing the dynamic mean process in such the datasets, our focus is on analyzing the tail-dependence between the stock prices of Indigo and SpiceJet. Understanding the stochastic behavior in both the tails is important, as buy positions are sensitive to falling prices (lower tail), while sell positions are sensitive to the upper tail prices. In order to assess both types of risk, we will use a recently proposed copula model which is able to distinctively capture asymptotic dependence or independence in its lower and upper tail simultaneously.

We here apply our model to the log return data in average price data of two big airlines in India, Indigo and SpiceJet Airlines and try to infer their tail behaviour. We also did simulation studies to assess our model before feeding in real data. From a practical perspective, our results would help detecting the market risk contagion and would be a potential source of information for investors who seek to diversify their portfolio.

1 Introduction

The year 2020 has been hit severely by a pandemic of a new kind, caused by the severe acute respiratory syndrome coronavirus 2 (SARS COV 2), a new strain of coronavirus that had not previously been identified in humans. The new strain was named COVID-19 on February 11, 2020 by WHO. On 11th March 2020, the WHO director announced the outbreak of COVID-19 as a pandemic due to the rapid increase in the number of cases outside China. Learning from the behaviour of virus it was found out that SARS-CoV-2 virus mainly spreads through cough droplets and is contagious. In view of this most countries adopted partial or complete Lockdown to prevent the spread of disease. Such measures have helped contain the pandemic to some extent. However, an immediate consequence was the effect on economy. International travelling was also affected since many country governments restricted the movement across boundaries.

Historically it has been seen that transportation industry is subjected to system risks which can be triggered by external factors like natural calamity or financial events etc. COVID pandemic is clearly one such event that which can highly increase the risk to airline sector as tourism has taken halt in most countries. Fear of a global recession, the year 2020 challenged the traditional financial system. For investors, behaviour of both tail is important since buy positions are risky in falling prices and sell positions are risky in price hike.

In each tail, there can be two type of dependency class asymptotic dependence or asymptotic independence. In this paper, we will see a single flexible dependence model for the entire dataset that possesses high flexibility in both the lower and the upper tails and can transit between the two. We here apply our model to the log return data in average price data of two big airlines in India, Indigo and SpiceJet Airlines and try to infer their tail behaviour

2 Definitions

Tail Dependence

The tail dependence of a pair of random variables (X_1, X_2 say) is a measure of their co-movements in the tails of the distributions i.e. It is conditional probability of random variable V_1 taking an extreme value given that I already know that V_2 has taken extreme value.
Consider Example - Tail dependence: basically refers to the conditional probability of an extreme move happening in one variable in V_2 given that I already know that V_1 has already moved V_2. If you were to take a look at two proper distributions that we use,
lets say a normal distribution vs lets say lets student t distribution. Then we will find that stu-

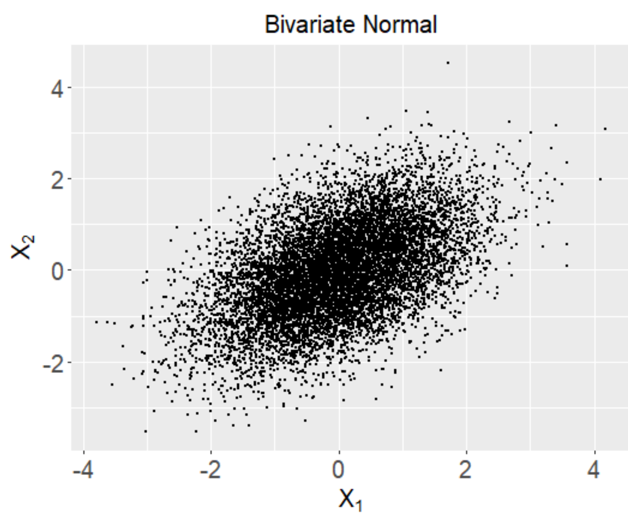


Figure 1: Scatter Plot of Random variables with Normal Distribution

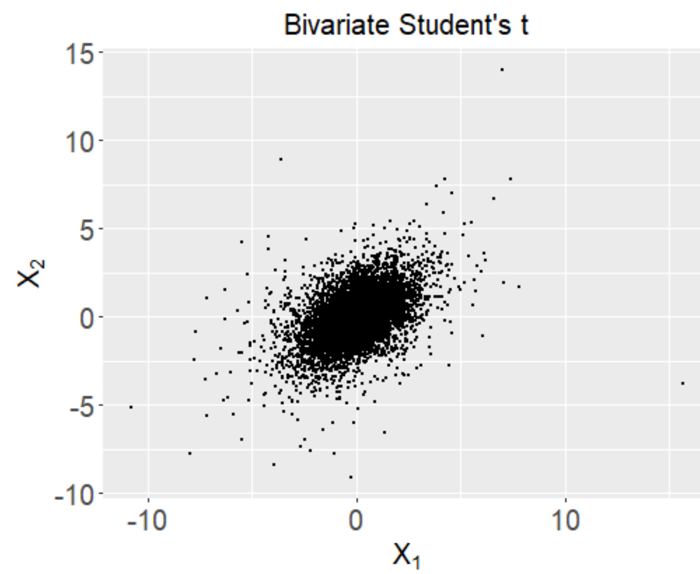


Figure 2: Scatter Plot of Random variables with t Distribution

dent's t distribution has more tail dependence as compared to normal distribution which means that it is more common in t distribution for the two variables v1 and v2 to be jointly experiencing extreme moves, as compared to if these two variables were normally distributed. That is the if these two were variables were students t distributed there is a higher chance that these two distributions were found together in the tail. it can be the right tail and it can be the left tail and this chance of them together in being the tail is higher as compared to normal . Let's take a look at Figure 1 and 2. Now using above explanation one can infer that Figure 2 one is student t distribution and Figure 1 is normal distribution.

Copula

In probability theory and statistics, a copula is a multivariate cumulative distribution function for which the marginal probability distribution of each variable is uniform on the interval [0, 1]. Copulas are used to describe/model the dependence (inter-correlation) between random variables. Copula can be used to find joint distribution of random variables whose marginals are known, since their joint distribution may not be easy to calculate. Mathematically, let X_1 and X_2 be two random variables with marginals F_{X_1} and F_{X_2}, and define the Uniform random variables $U_1 = F_{X_1}(X_1)$, $U_2 = F_{X_2}(X_2) \sim Unif(0,1)$. Then, joint distribution of vector $\mathbf{U} = (U_1, U_2)^\top$ is defined by copula as,

$$C(u_1, u_2) = F_{\mathbf{X}}\{F_{X_1}-1(u_1), F_{X_2}-1(u_2)\},$$

Here $C(u_1, u_2)$ is called the copula of $\mathbf{X}$.

Asymptotic Dependence/Independence

Asymptotic dependence can be interpreted as the property that realization of random vectors occur simultaneously with extreme values. Mathematically, lower tail dependence of random vector $\mathbf{X}$ is determined by tail coefficients,

$$\chi_L(t) = Pr(U_1 < t | U_2 < t) = \frac{C(t,t)}{t}, \quad t \in (0,1)$$

where $\chi_L(t)$ is lower tail dependence coefficients. So a random vector is said to be AD in lower tail if $\lim_{t\to 0} \chi_L(t) > 0$ and AI if limit is 0. Similarly, upper tail asymptotic dependence can be determined by taking limit on upper tail coefficient $\chi_U(t)$, we will discuss them later.

Stationary Series

A stationary time series is one whose properties do not depend on the time at which the series is observed. Thus, time series with trends, or with seasonality, are not stationary — the trend and seasonality will affect the value of the time series at different times. In general, a stationary time series will have no predictable patterns in the long-term. Time plots will show the series to be roughly horizontal (although some cyclic behaviour is possible), with constant variance.
As well as looking at the time plot of the data, the ACF plot is also useful for identifying non-stationary time series. For a stationary time series, the ACF will drop to zero relatively quickly, while the ACF(auto-correlation function) of non-stationary data decreases slowly. Also, for non-stationary data, the value of r1 is often large and positive.

3 Analysing Data set and Data Preprocessing

Airline is a vital, but brutal industry. Many have entered this space, but very few have survived. The bottom line is, the industry which was already gasping for breath, has been forced on its knees by the Covid-19 pandemic. Inter globe Aviation, which runs low cost carrier Indigo, for the June quarter posted losses at the rate of Rs 35 crore daily, which was larger than any expectation. This loss took its combined loss for the last six quarters to close to Rs 10,000 crore. The company said its business suffered heavily due to the second wave of the pandemic.(Source: The Economic Times) Spice Jet also stands in the same line. It reported a net loss of Rs 600.5 crore for the first quarter ended June 30, owing to the suspension of flight operations due to the coronavirus-induced lockdown.(Source: business-standard.com). Also there can many such situations when log returns of both stock prices may get affected. Thus, this study is devoted to analyse the log return of

stock price data of two major Airlines in India.
For this study, historical prices stock price data of airlines can be downloaded from Investing.com. Figure 3 shows Daily average price and log return of two major Indian Airlines and Spice Jet from March 11, 2020 to March 25, 2022.

In order to check stationarity of data We also plotted the ACF plot. Figure 4(a) shows the ACF plot of Indigo and Spice Jet log return data. From plot it can inferred that our data is pretty much stationary. Partial Auto-correlation Function is also plotted as shown in Figure 4(b) to further check the correlation between stationary data. Figure 5(b) displays the bivariate scatter plots of log return data of Indigo and Spice Jet and Figure 5(b) shows the bivariate plot of data after rank transformation. Both Figure shows that there is dependence in their average returns and also that the strength if dependence is more in upper tail than lower tail.

To further check the dependence we calculated empirical upper-upper, lower-upper, upper-lower and lower-lower tail dependence coefficients by taking 0.95 levels using following R code -

```
mean((Indigo.rank > 0.95) & (SpiceJet.rank > 0.95)) / 0.05
mean((Indigo.rank > 0.95) & (SpiceJet.rank < 0.05)) / 0.05
mean((Indigo.rank < 0.05) & (SpiceJet.rank > 0.95)) / 0.05
mean((Indigo.rank < 0.05) & (SpiceJet.rank < 0.05)) / 0.05
```

here Indigo.rank and Spicejet.rank is rank transformation of our corresponding log return data and got the following output. Above code gave following output.

Upper-Upper - 0.48484848
Upper-Lower - 0.04040404
Lower-Upper - 0.00000000
Lower-Lower - 0.24242424

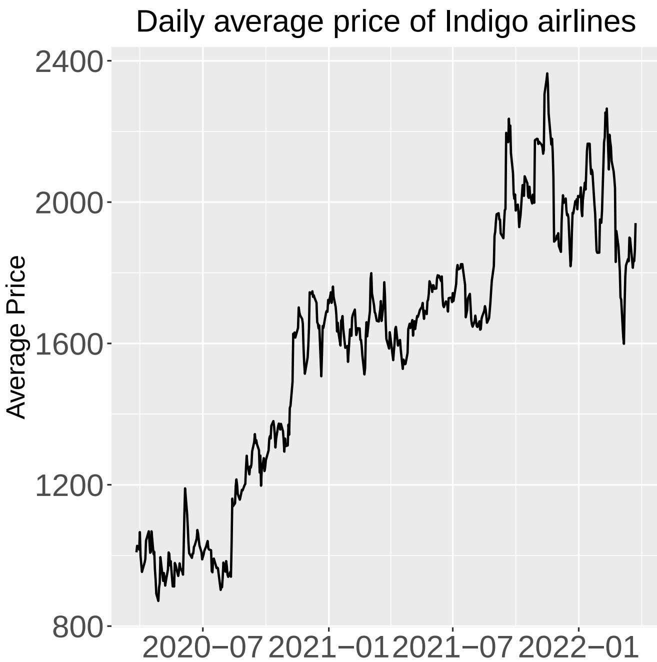
To further check the dependence we calculated empirical upper-upper, lower-upper, upper-lower and lower-lower tail dependence coefficients by taking 0.98 levels also, following is result we obtained

Upper-Upper - 0.4040404
Upper-Lower - 0.0000000
Lower-Upper - 0.0000000
Lower-Lower - 0.0000000

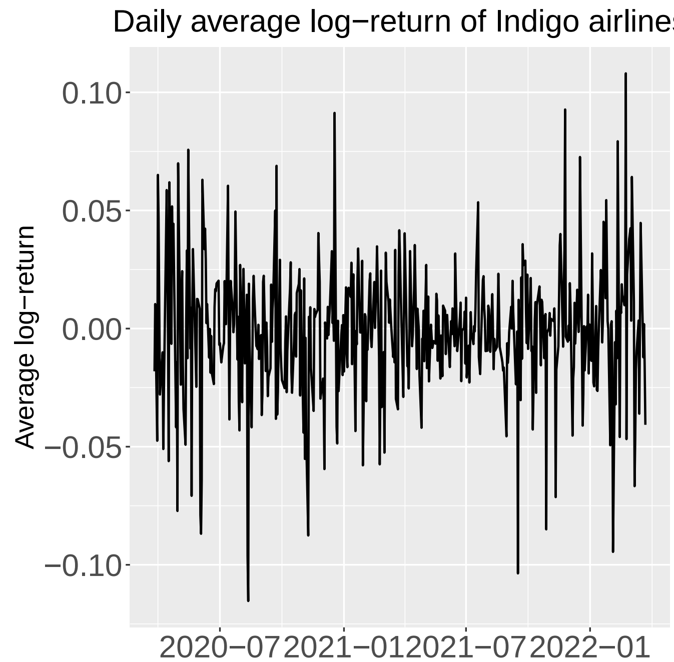
Thus all tail dependencies except upper-upper seems to vanish asymptotically. We can infer from this output data that there more upper-upper tail dependence than lower-lower, upper-lower or lower-upper, that it is highly likely that if return from one stock is higher then it would same for another stock. Also we see that stock there is very less dependence in lower-upper, that is if there is low return from Indigo stocks then it is very unlikely to get high returns from spicejet stock.
We have seen enough evidences that there is some dependence between the two airline log return data and also that upper tail dependence is more than other tail dependencies.

4 Modeling

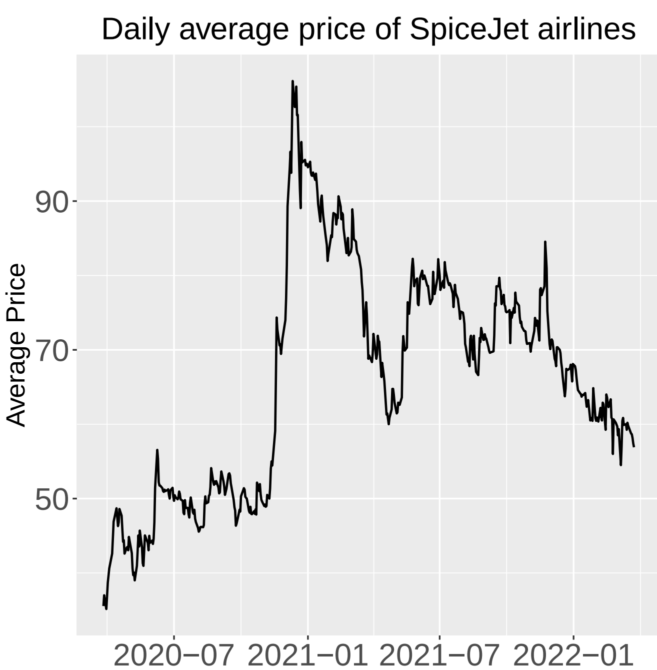
So far we have seen empirically that there is dependence in returns of both airlines. Now we will see a flexible copula model which is able to distinctively capture asymptotic dependence or independence in its lower and upper tails simultaneously. This model is parsimonious and smoothly bridges (in each tail) both extreme dependence classes in the interior of the parameter space



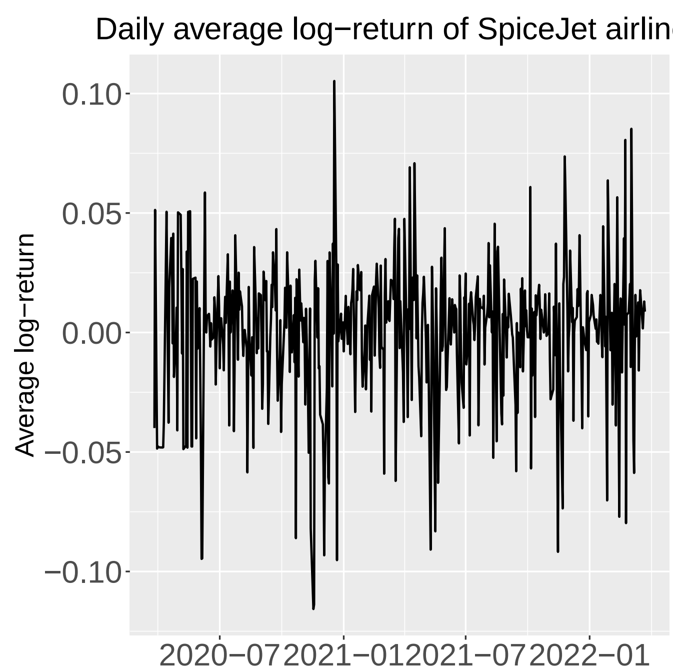
(a)



(b)

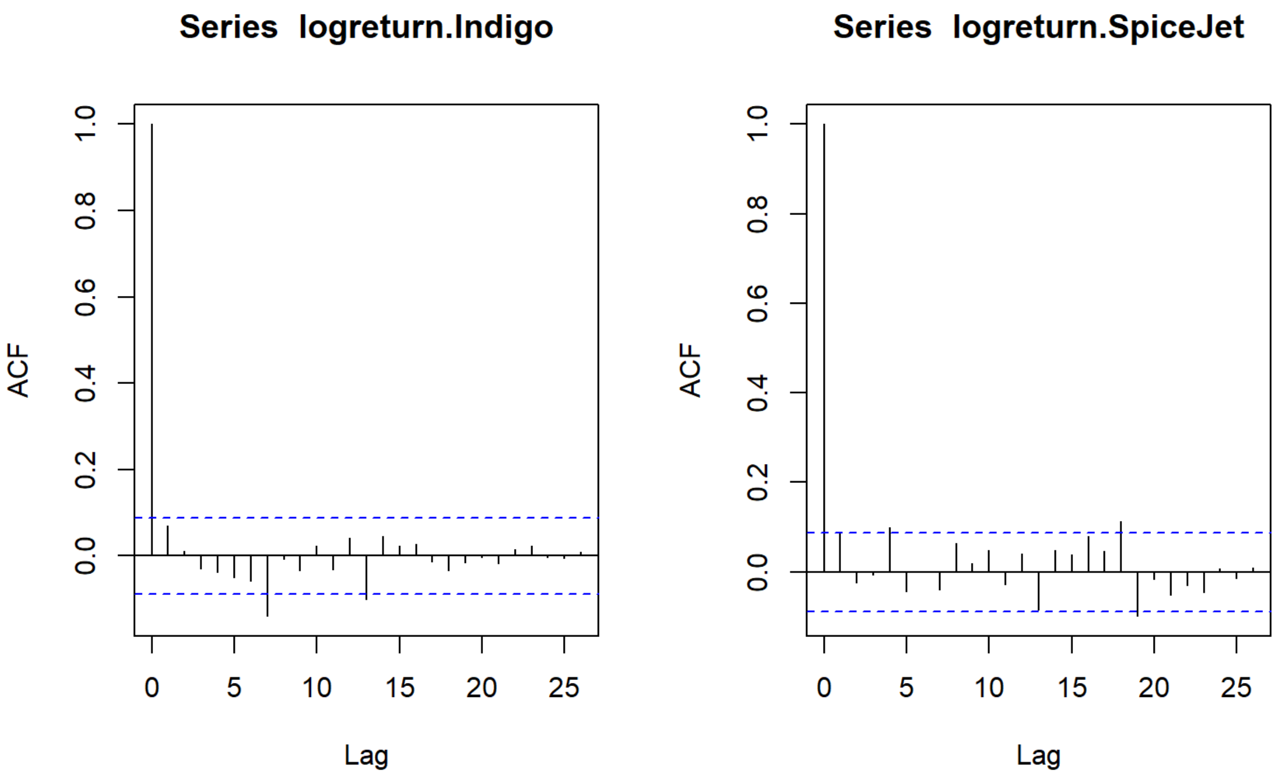


(c)

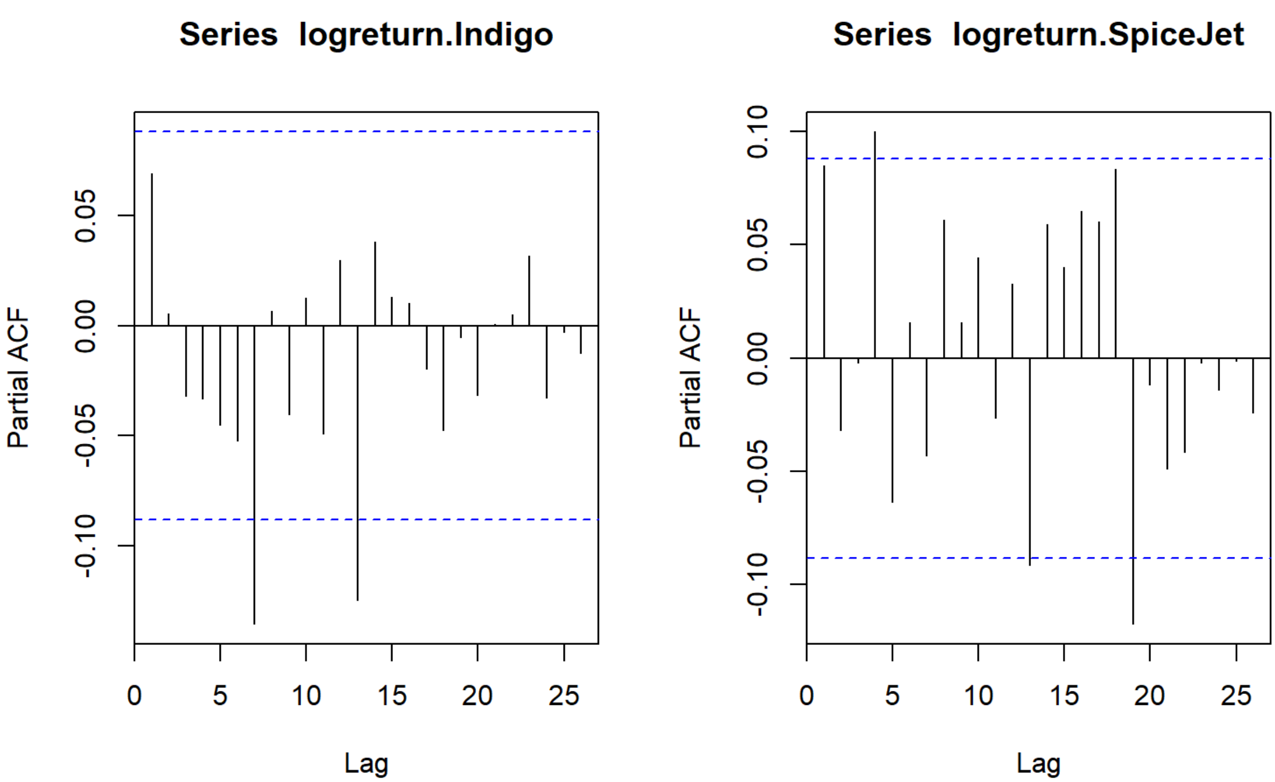


(d)

Figure 3: Daily average price and log-return plots of Indigo and SpiceJet airlines taken from Mar 11, 2020 to Mar 25, 2022



(a) ACF plot of Indigo (left) and SpiceJet (right) log return data



(b) PACF plot of Indigo and SpiceJet log return data

Figure 4: Auto-correlation (ACF) and Partial auto-correlation plots to check stationarity of data

(a) Bivatriate Scatter Plot of Indigo and SpiceJet of log return data after rank transformation

(b) Bivariate Sctter plot of Standardised log return data of Indigo and SpiceJet Airline

Figure 5: It can inferred that there is more upper-upper tail dependence as compared to other tail dependencies

Model Construction

Here we will construct the model to capture lower and upper tail dependence, Let R be random variable with asymmetric Laplace distribution, say F_R given by $AL(\lambda_L, \lambda_U)$,

$$F_R(r) = \begin{cases} \frac{\lambda_L}{\lambda_L+\lambda_U} exp(r/\lambda_L), & r \leq 0, \\ 1 - \frac{\lambda_U}{\lambda_L+\lambda_U} exp(-r/\lambda_U), & r > 0, \end{cases} r \in \mathbb{R}$$

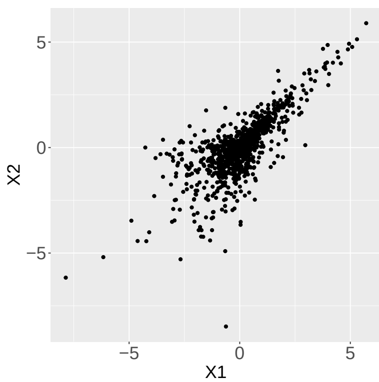
where $\lambda_L, \lambda_U \in (0,1)$ are scale parameters of lower and upper tails, respectively. Also, Let W_1, W_2 have $AL(1-\lambda_L, 1-\lambda_U)$ distribution, then wee assume that the random vector $\mathbf{W} = (W_1, W_2)^\top$ follows Gaussian Copula model with correlation $\rho \in (-1, 1)$, i.e. joint distribution of $\mathbf{W}$ satisfies,

$$Pr(W_1 \leq w_1, W_2 \leq w_2) = \Phi_\rho[\Phi^{-1}\{F_W(w_1)\}, \Phi^{-1}\{F_W(w_2)\}],$$

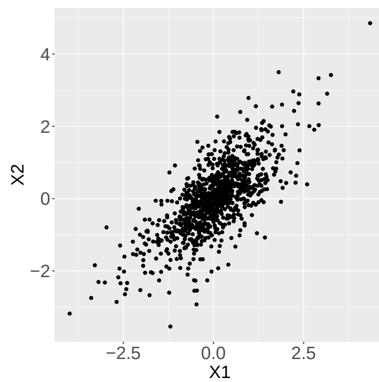
where Φ and Φ_ρ denote the uni-variate standard Gaussian distribution and bi-variate standard Gaussian distribution with correlation ρ, respectively. We define our model through the random vector $\mathbf{X} = (X_1, X_2)^\top$ with components

$$X_1 = R + W_1, \quad X_2 = R + W_2.$$

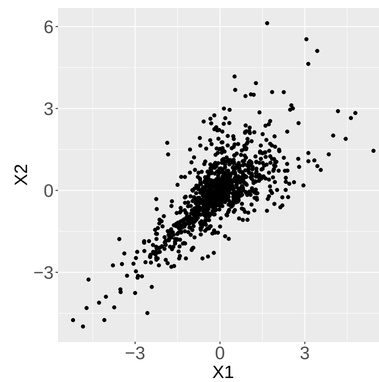
Note that, random vector $\mathbf{X}$ is combination of random vector R and random vector $\mathbf{W}$, thus providing perfectly dependent structure and Gaussian dependence structure i.e asymptotically independence structure. Thus dependence structure of $\mathbf{X}$ interpolates between that $\mathbf{R}$(perfect dependence) and that of $\mathbf{W}$(independence).
Thus, our above model consists of three-parameters $(\lambda_L, \lambda_U, \rho)$ and dependence structure in both lower and upper tail can then be inferred by looking at these parameters only. To illustrate this,



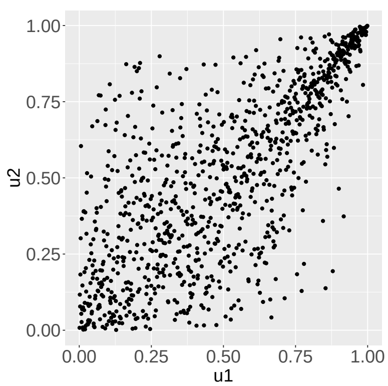
(a) $\lambda_L = 0.1,\ \lambda_U = 0.9$



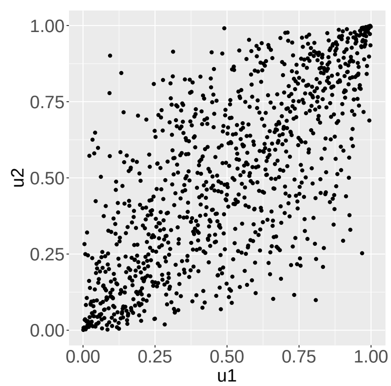
(b) $\lambda_L = 0.5,\ \lambda_U = 0.5$



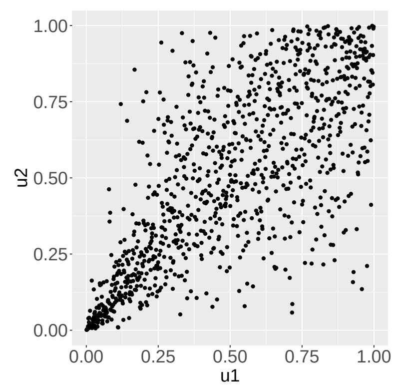
(c) $\lambda_L = 0.9,\ \lambda_U = 0.1$



(d) $\lambda_L = 0.1,\ \lambda_U = 0.9$



(e) $\lambda_L = 0.5,\ \lambda_U = 0.5$



(f) $\lambda_L = 0.9,\ \lambda_U = 0.1$

Figure 6: Bivariate Scatter plot and plot after rank transformation of random samples from our model for different values of λ_L, λ_U and $\rho = 0.5$. This Figure shows how model can bridge between different dependence regimes just by changing internal model parameters.

Figure 6. shows bivariate scatter plots, before and after rank transformation, obtained by sampling from our model with different parameter values.
Note that we might have used any other model, other than Gaussian copula model, that is asymptotically independent in both tails but we have chosen Gaussian copula model only for computational convenience.

Expressions for the associated copula

Now we state all the expressions of marginals and joint distributions deduced using above model and used in our source code to analyse real life data.
Here also we use same notations, Let f_R denote the asymmetric Laplace distribution $AL(\lambda_L, \lambda_U)$ obtained by differentiating F_R with respect to, r,

$$f_R(r) = \begin{cases} \frac{1}{\lambda_L+\lambda_U} exp(r/\lambda_L), & r \leq 0, \\ \frac{1}{\lambda_L+\lambda_U} exp(-r/\lambda_U), & r > 0, \end{cases} \quad r \in \mathbb{R}$$

Then Marginal distribution for X_1 and X_2 is,

$$
\begin{aligned}
F_X(x) =& Pr(X_i \le x) = Pr(R + W_i \le x) = \int_{\mathbb{R}} Pr(W_i \le x - r) f_R(r) dr \\
&= \frac{1}{\lambda_L + \lambda_U} \{ \int_{-\inf}^{0} Pr(W_i \le x - r) exp(r/\lambda_L) dr + \int_{0}^{\inf} Pr(W_i \le x - r) exp(-r/\lambda_U) dr \}
\end{aligned}
$$

Plugging all expressions in above equation we get marginal distribution of X_i for $\lambda_L, \lambda_U \neq 1/2$, as

$$
F_X(x) = \begin{cases} K_1(\lambda_L, \lambda_U) exp(\frac{x}{\lambda_L}) - K_2(\lambda_L, \lambda_U) exp(\frac{x}{1-\lambda_L}), & x \le 0, \\ 1 + K_3(\lambda_L, \lambda_U) exp(\frac{-x}{\lambda_U}) - K_4(\lambda_L, \lambda_U) exp(\frac{-x}{1-\lambda_U}), & x > 0, \end{cases} r \in \mathbb{R}
$$

where K_i are constant coefficients depending on $\lambda_L and \lambda_U$ only, there expression are given as,

$$
\begin{aligned}
K_1(\lambda_L, \lambda_U) &= \lambda_L^3 \{ (\lambda_L + \lambda_U)(2\lambda_L - 1)(1 + \lambda_L - \lambda_U) \}^{-1} \\
K_2(\lambda_L, \lambda_U) &= (\lambda_L - 1)^3 \{ (2\lambda L - 1)(\lambda_L - \lambda_U - 1)(2 - \lambda_L - \lambda_U) \}^{-1} \\
K_3(\lambda_L, \lambda_U) &= \lambda_U^3 \{ (\lambda_L + \lambda_U)(2\lambda_L - 1)(\lambda_L - \lambda_U - 1) \}^{-1} \\
K_4(\lambda_L, \lambda_U) &= (\lambda_U - 1)^3 \{ (1 + \lambda_L - \lambda_U)(2\lambda_L - 1)(2 - \lambda_L - \lambda_U) \}^{-1}
\end{aligned}
$$

We can see above distribution diverge for $\lambda_L or \lambda = 1/2$. The cases when $\lambda_L = 1/2$ and/or $\lambda_U = 1/2$ can be calculated taking limits $\lambda_L \to 1/2$ and/or $\lambda_U \to 1/2$ When $\lambda_L = 1/2$ and $\lambda_U \neq 1/2$

$$
F_X(x) = \begin{cases} K_5(\lambda_L, \lambda_U) x exp(2x) - K_6(\lambda_L, \lambda_U) exp(2x), & x \le 0, \\ 1 - K_7(\lambda_L, \lambda_U) exp(\frac{x}{\lambda_L}) - K_8(\lambda_L, \lambda_U) exp(\frac{-x}{1-\lambda_U}), & x > 0, \end{cases}
$$

and expressions of normalising coefficients in above definition are given as

$$
\begin{aligned}
K_5(\lambda_L, \lambda_U) &= 2\{ (1 + 2\lambda_U)(2\lambda_L - 3) \}^{-1} \\
K_6(\lambda_L, \lambda_U) &= (-12\lambda_U + 12\lambda_U^2 - 5)\{ (1 + 2\lambda_U)^2 (2\lambda_L - 3)^2 \}^{-1} \\
K_7(\lambda_L, \lambda_U) &= 4\lambda_L^3 \{ (1 + 2\lambda_U)^2 (2\lambda_U - 1) \}^{-1} \\
K_8(\lambda_L, \lambda_U) &= 4(\lambda_U - 1)^3 \{ (2\lambda_U - 3)^2 (2\lambda_U - 1) \}^{-1}
\end{aligned}
$$

When $\lambda_L \neq 1/2$ and $\lambda_U = 1/2$

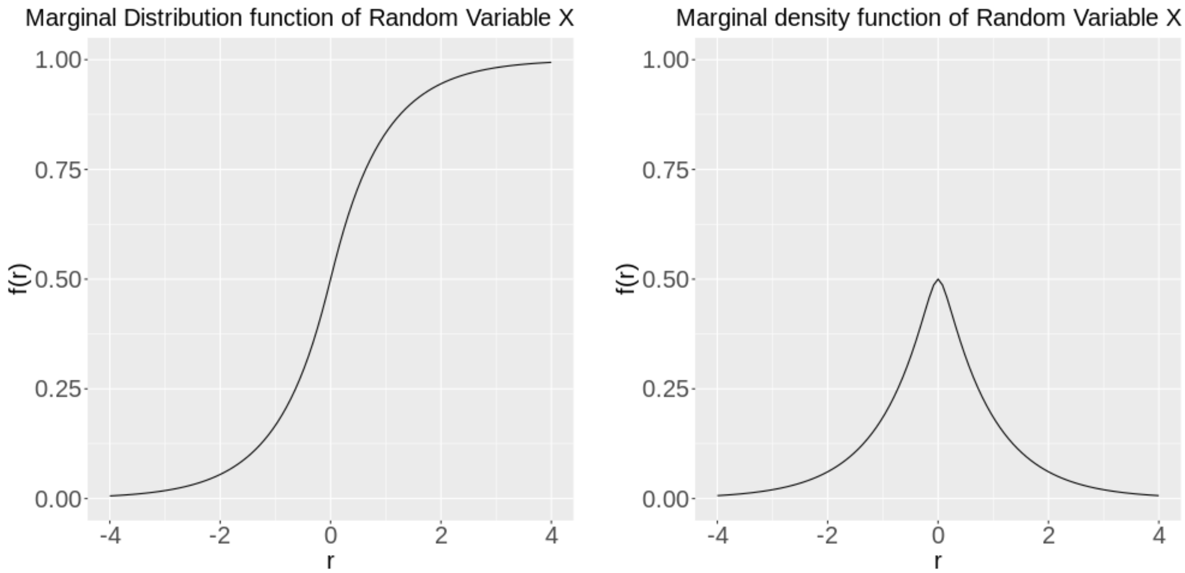
$$
F_X(x) = \begin{cases} K_9(\lambda_L, \lambda_U) exp(\frac{x}{\lambda_L}) + K_{10}(\lambda_L, \lambda_U) exp(\frac{x}{1-\lambda_L}), & x \le 0, \\ 1 + K_{11}(\lambda_L, \lambda_U) x exp(-2x) + K_{12}(\lambda_L, \lambda_U) exp(-2x), & x > 0, \end{cases}
$$

and expressions of normalising coefficients in above definition are given as

$$
\begin{aligned}
K_9(\lambda_L, \lambda_U) &= 4\lambda_L^3 \{ (1 + 2\lambda_L)^2 (2\lambda_L - 1) \}^{-1} \\
K_{10}(\lambda_L, \lambda_U) &= 4(\lambda_U - 1)^3 \{ (2\lambda_L - 3)^2 (2\lambda_L - 1) \}^{-1} \\
K_{11}(\lambda_L, \lambda_U) &= 2\{ (1 + 2\lambda_L)^2 (2\lambda_U - 3) \}^{-1} \\
K_{12}(\lambda_L, \lambda_U) &= 4(\lambda_U - 1)^3 \{ (2\lambda_L - 3)^2 (1 + 2\lambda_L) \}^{-1}
\end{aligned}
$$

When $\lambda_U = \lambda_L = 1/2$,

$$
F_X(x) = \begin{cases} \frac{1}{2}(1 - x) exp(2x), & x \le 0, \\ 1 - \frac{1}{2}(1 + x) exp(-2x), & x > 0, \end{cases}
$$



(a) (b)

Figure 7: Marginal distribution F_X and density plot f_X of random variable X for $\lambda_L = 0.9,\ \lambda_U = 0.9$

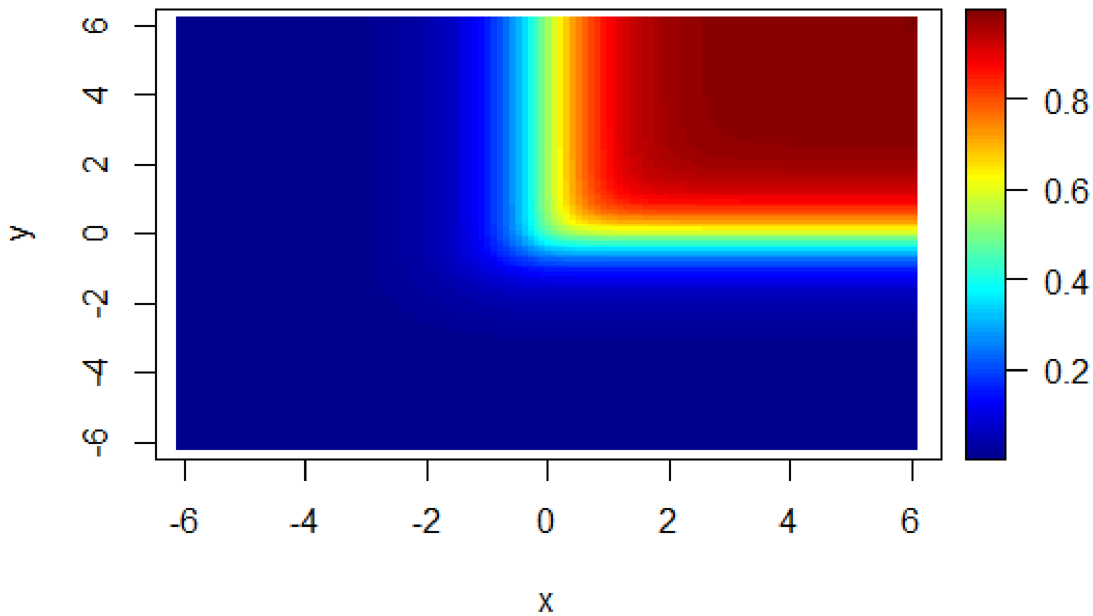
Figure 7 shows the plot of our distribution and density function of random variable X_1 to get better visualisation. Since we have explicit expression of distribution of $\mathbf{X}$, we can obtain it density function f_X easily by differentiating F_X. And joint distribution $F_X(x_1, x_2)$ can be obtained as

$$\begin{aligned} F_X(x_1, x_2) &= Pr(X_1 \leq x_1, X_2 \leq x_2) \\ &= Pr(R + W_1 \leq x_1, R + W_2 \leq x_2) \\ &= \int_{\mathbb{R}} \Phi_\rho[\Phi^{-1}\{F_W(x_1 - r)\}, \Phi^{-1}\{F_W(x_2 - r)\}] f_R(r) dr \end{aligned}$$

where Φ^{-1} is quantile function of standard normal distribution and Φ_ρ is standard Normal distribution with correlation $\rho \in (-1, 1)$, and F_W is $AL(1 - \lambda_L, 1 - \lambda_U)$. Since we don't have explicit expression for quantile of Gaussian distribution, thereby we have to use numerical integration for implementation of distribution function. Now to get joint density function we need to differentiate above expression with respect to x_1 and x_2, Figure 8 shows how joint distribution will look like.

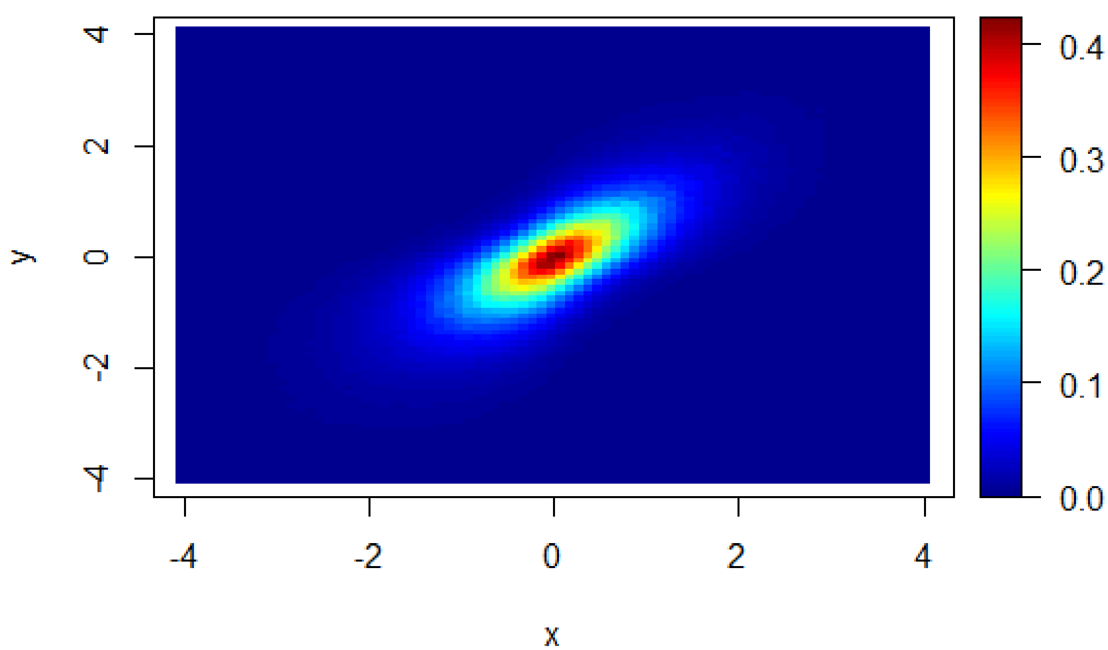
$$\begin{aligned} f_X(x_1, x_2) &= \int_{\mathbb{R}} \frac{\partial^2 f}{\partial x_1 \partial x_2} Pr(W_1 \leq x_1 - r, W_2 \leq x_2 - r) f_r(r) dr \\ &= \int_{\mathbb{R}} \phi_\rho[\Phi^{-1}\{F_W(x_1 - r)\}, \Phi^{-1}\{F_W(x_2 - r)\}] \prod_{i=1}^{2} \frac{f_W(x_i - r)}{\phi\{F_W(x_i - r)\}} f_R(r) dr \end{aligned}$$

here $\phi,\ \phi_\rho$ denotes univarite standard Gaussian density and bivariate Gaussian density with correlation ρ, respectively. Figure 9 shows the plot of joint density function.



(a)

Figure 8: Joint distribution $F_{\mathbf{X}}$ plot of random vector $\mathbf{X}$ for $\lambda_L = 0.3,\ \lambda_U = 0.4$ and $\rho = 0.7$



(a)

Figure 9: Joint density $f_{\mathbf{X}}$ plot of random vector $\mathbf{X}$ for $\lambda_L = 0.3,\ \lambda_U = 0.4$ and $\rho = 0.7$

Tail Dependence Coefficients

We discussed earlier that, tail dependence can be determined by using coefficients $\chi_L(t)\ and\ \chi_U(t)$. Here t is threshold value, i.e. for $U_1 \sim F_R$, $U_1 > t$ implies extreme value. Thus t can take value in $(0,1)$ Let $U_1 \sim F_{X_1}\ and\ U_2 \sim F_{X_2}$, and by definition

$$C(U_1, U_2) = F_{\mathbf{X}}\{F_{X_1}-1(U_1), F_{X_2}-1(U_2)\},$$

we can calculate $\chi_L(t)\ and\ \chi_U(t)$, as

$$\begin{aligned}\chi_L(t) &= Pr(U_1 < t|U_2 < t) \\ &= \frac{Pr(0 < U_1 < t, 0 < U_2 < t)}{Pr(U_2 < t)} \\ &= \frac{C(t,t) + C(0,0) - C(0,t) - C(t,0))}{t} \\ &= \frac{C(t,t)}{t}\end{aligned}$$

Similarly we can calculate,

$$\begin{aligned}\chi_U(t) &= Pr(U_1 > t|U_2 > t) \\ &= \frac{Pr(t < U_1 < 1, t < U_2 < 1)}{Pr(t < U_2 < 1)} \\ &= \frac{1 - 2t - C(t,t)}{1 - t}\end{aligned}$$

and Asymptotic dependence class (AD/AI) can be determined by computing limits $\lim_{t\to 0} \chi_L(t) > 0$ and $\lim_{t\to 0} \chi_U(t) > 0$. In order to visualize the various types of dependence structures that our model can produce , Figure 10 displays $\chi_L(t)$ and χ_U for t $\in$ (0, 1) for different parameter values.

High fluctuation at edges in above plot may be due numerical instability near boundary values.

5 Estimating parameters

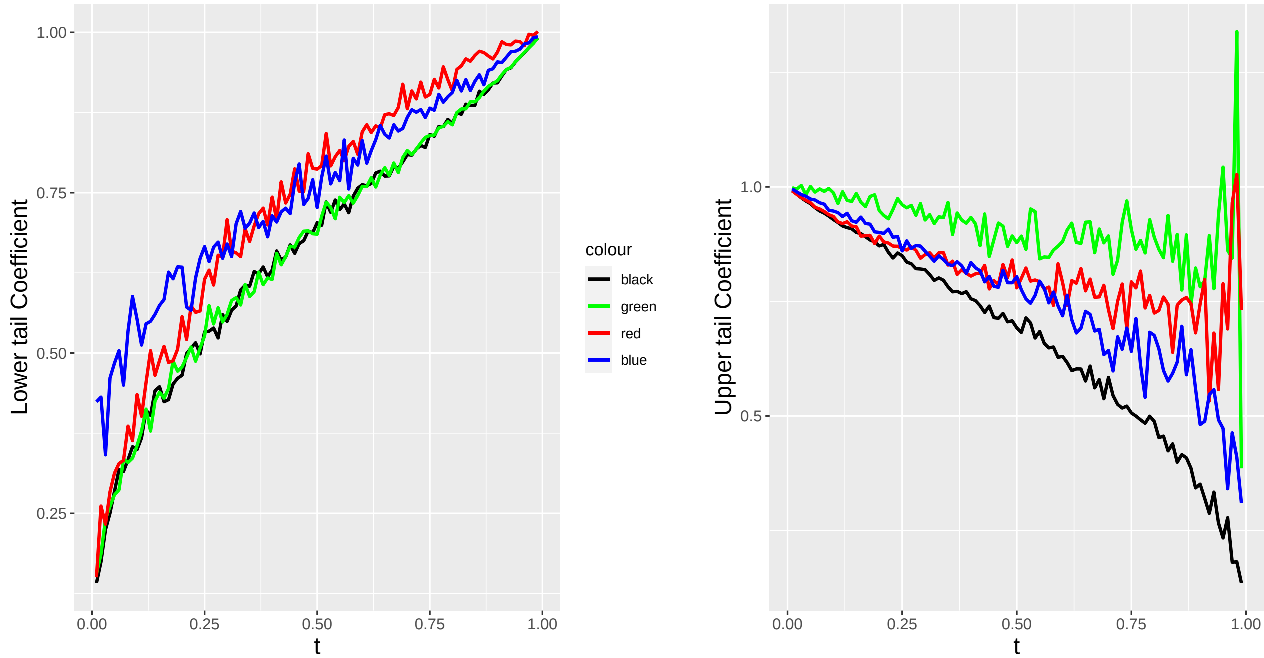
To estimate dependence parameters we use maximum likelihood approach. Likelihood for our copula model may be written as,

$$\begin{aligned}L(\theta) &= \prod_{i=1}^{n} c(u_{i1}, u_{i2}), \\ \theta &= (\lambda_L, \lambda_U, \rho)^\top \in (0,1) \times (0,1) \times (-1,1)\end{aligned}$$

where $c(u_1, u_2)$ is copula density defined as

$$c(u_{i1}, u_{i2}) = \frac{f_X\{F_X^{-1}(u_1), F_X^{-1}(u_2)\}}{f_X\{F_X^{-1}(u_1)\} f_X\{F_X^{-1}(u_2)\}}$$

it depends on model parameters $\theta = (\lambda_L, \lambda_U, \rho)^\top$. Maximising $L(\theta)$ will give us maximum likelihood estimator $\hat{\theta}$ and thereby we can analyse the dependence structure of our data.



(a) Lower tail coefficients plot for different values of λ_L and λ_U

(b) Upper tail coefficients plot for different values of λ_L and λ_U

Figure 10: Coefficients with threshold t $\in [0.01, 0.99]$, with correlation $\rho = 0.5$ and tail parameters $\lambda_L = 0.3,\ 0.8,\ 0.3,\ 0.5$ (black,green, red, blue), $\lambda_U = 0.2,\ 0.8,\ 0.8,\ 0.5$ respectively

6 Simulation study

To check the performance of our model we now conduct simulation study by fixing the parameters and then simulating data from our model then again feed it to likelihood function and let it estimate parameters.

Data Simulation

To simulate random vector $\mathbf{X} = (X_1, X_2)$, in our model we simulate $\mathbf{R}$= (R,R) and $\mathbf{W}$= (W_1, W_2) separately and then simply add them since

$$X_i = R + W_i$$

$R \sim AL(\lambda_L, \lambda_U)$, To simulate R we first simulate standard uniform variables and then feed them to quantile function of Asymmetric Laplace given by

$$F_R^{-1}(p) = \begin{cases} \lambda_L(log(p) + log(\frac{\lambda_L+\lambda_U}{\lambda_L}) & p \leq \frac{\lambda_L}{\lambda_L+\lambda_U}, \\ -\lambda_U(log(1-p) + log(\frac{\lambda_L+\lambda_U}{\lambda_U}) & p > \frac{\lambda_L}{\lambda_L+\lambda_U} \end{cases}$$

$W_1, W_2 \sim F_W$ have the $AL(1-\lambda_L, 1-\lambda_U)$ distribution, and bivariate random vector $\mathbf{W} = (W_1, W_1)$ is driven by a Gaussian copula with correlation $\rho \in (-1, 1)$, To simulate $\mathbf{W}$, we first simulate random vector $\mathbf{N}$ from bivariate standard normal distribution with correlation ρ, then compute $Y_i = \Phi(N_i)$, then feed each component to quantile function of $AL(1-\lambda_L, 1-\lambda_U)$ to get W_i

$$\begin{aligned} &(N_1, N_2) \sim \mathcal{N}(0, \Sigma) \\ &W_1 = F_W^{-1}(\Phi(N_1)), W_2 = F_W^{-1}(\Phi(N_2)) \end{aligned}$$

We start by simulating n = 500, 1000 independent samples from our model with correlation $\rho = 0.2$ and tail parameters $\lambda_L = 0.7$, $\lambda_U = 0.4$ (Case 1: strong lower tail dependence, weak upper tail dependence), $\lambda_L = 0.7$, $\lambda_U = 0.8$ (Case 2: strong lower and upper tail dependence), These cases covers various combinations of extremal tail dependence in each tail. And then we feed our data to likelihood to estimate our our model parameters $\theta = (\lambda_L, \lambda_U, \rho)$ We then repeat this experiment 10 times to produce box plots of estimated parameters. Figure 11 shows the box plots of parameters in Case 1 with 500 and 1000 independent samples.

7 Application on Ingl and Spicejet Dataset

We now apply our model to uncover the tail dependence of real log return data of Indigo and Spicejet airlines. In this section, we assume that the dependence structure is stationary after our data preprocessing. To get log return data in uniform scale rank transformation is applied and then feed to likelihood function which then optimised to get parameter values from which dependence structure of data can be inferred.
Thus fitting our model to log-return data of Indigo and SpiceJet airline we get the following values of parameters.

$$\lambda_L = 0.2556414$$
$$\lambda_U = 0.5027140$$
$$\rho = 0.2145668$$

that stands well with our empirical estimations, though there are some deviations from empirical estimations. We also calculated standard error in our parameters using hessian matrix H that we obtain in *optim* function while optimising the likelihood function for real data.

$$Standard\ error = \sqrt{diag(H^{-1})}$$

here diag means diagonal entries of inverse of hessian matrix and then square root of entry is taken. Thus result we obtained is as follows

$$Standard\ error\ in\ \lambda_L = 0.000595$$
$$Standard\ error\ in\ \lambda_U\ = 0.003074$$
$$Standard\ error\ in\ \rho = 0.003860$$

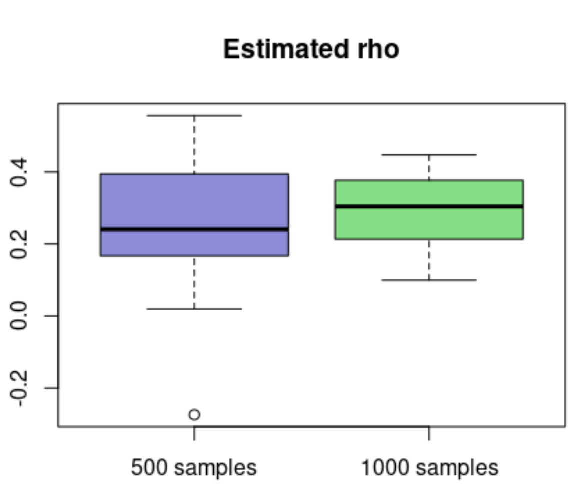
Also after estimating parameters we can estimate asymptotic regime of each tail. To visualise asymptotic tail behaviour we plot $\chi_L and \chi_U$ against threshold t using expression given in **Tail dependence Coefficients** section that is

$$\chi_L(t) = \frac{C(t,t)}{t}, \quad \chi_U(t) = \frac{1 - 2t - C(t,t)}{1 - t}$$

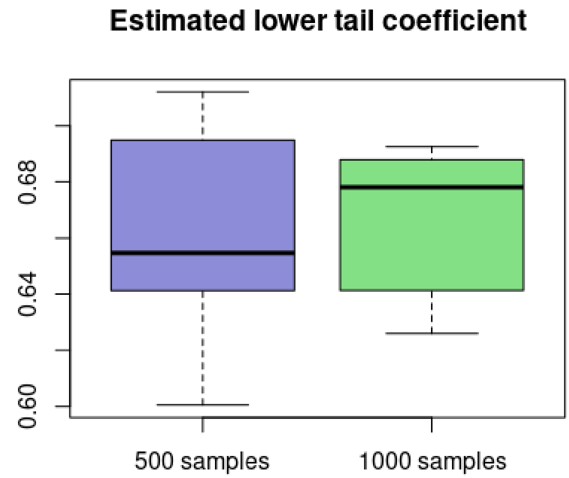
Figure 12 shows plots of dependence coefficients using estimated parameters obtained by fitting real data to our model. Thus we see

8 Conclusion

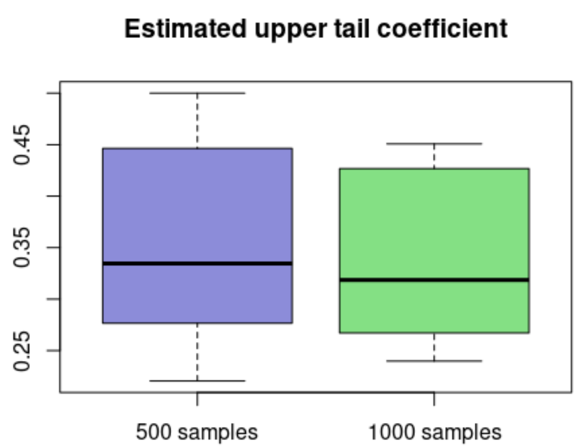
In this paper, we have explored a new parsimonious copula model that possesses high flexibility in both the lower and upper tails. This model can bridge from asymptotic dependence and



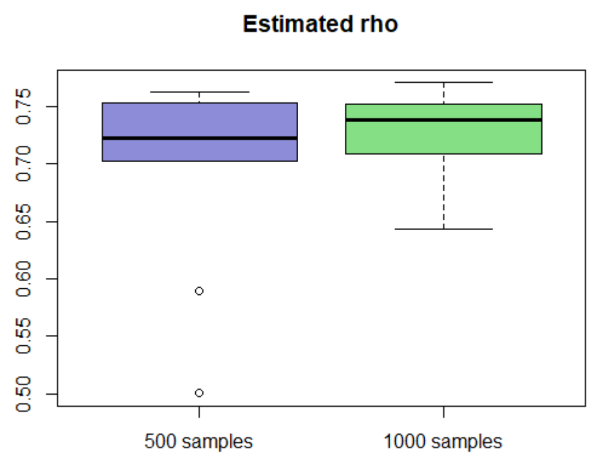
(a) True values $\lambda_L = 0.7$ $\lambda_U = 0.4$ and $\rho = 0.2$



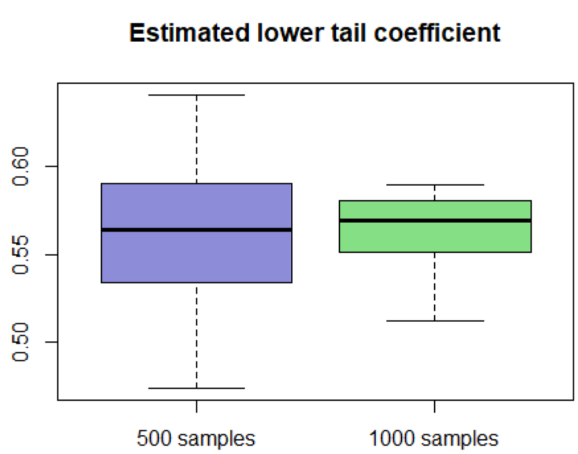
(b) True values $\lambda_L = 0.7$ $\lambda_U = 0.4$ and $\rho = 0.2$



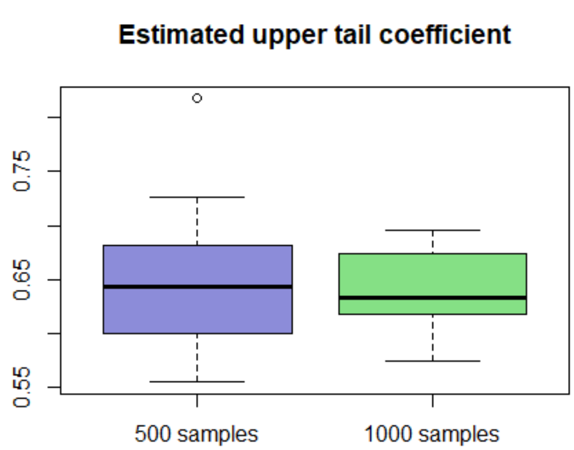
(c) True values $\lambda_L = 0.7$ $\lambda_U = 0.4$ and $\rho = 0.2$



(d) True values $\lambda_L = 0.7$ $\lambda_U = 0.4$ and $\rho = 0.2$

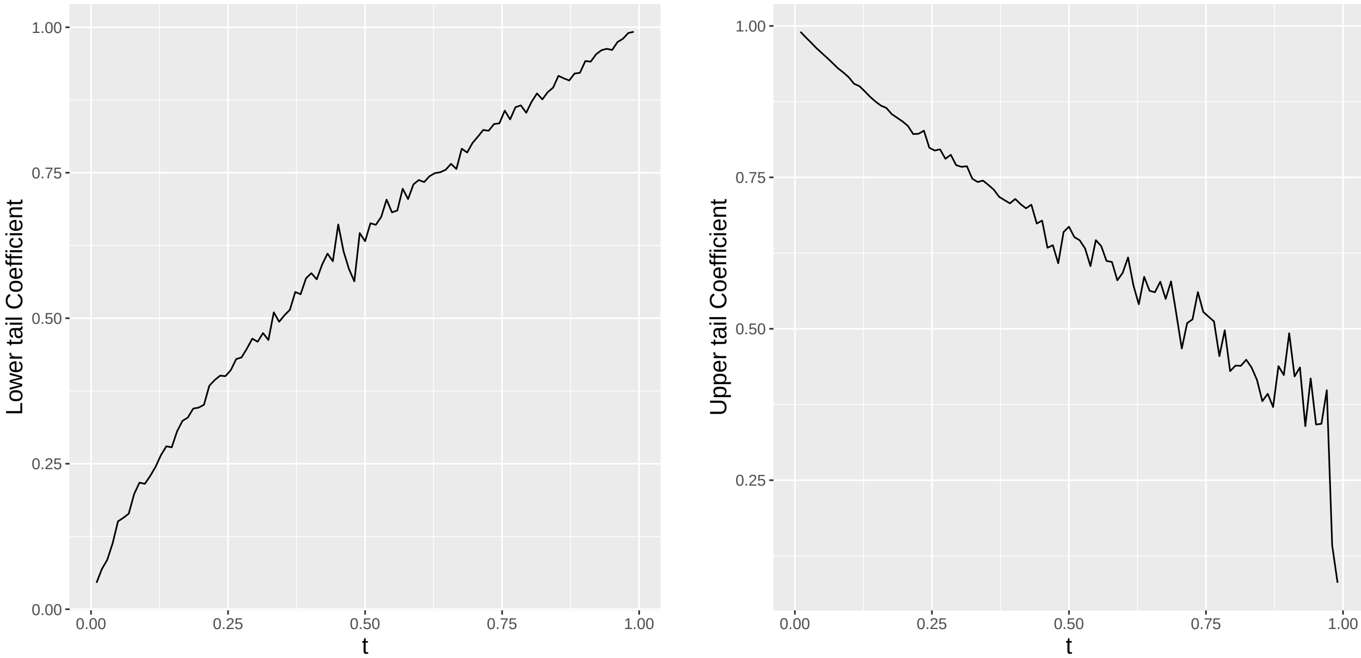


(e) True values $\lambda_L = 0.7$ $\lambda_U = 0.4$ and $\rho = 0.2$



(f) True values $\lambda_L = 0.7$ $\lambda_U = 0.4$ and $\rho = 0.2$

Figure 11: Box plots of estimated parameters of sample data with different true values of parameter and with sample size of n = 500 and 1000



(a) χ_L plot against threshold t

(b) χ_U plot against threshold t

Figure 12: χ plots against threshold t using estimated parameters obtained by fitting real data to our model

independence in the interior of the parameter space, which simplifies inference on the extremal dependence class that is we can determine the class of tail (AI/AD) looking at model parameter only. Inference is performed by maximum likelihood approach.
We also checked our model by fitting sample data obtained by sampling using true parameters and then estimated them again by fitting to likelihood function, and plotted the estimates as box plots (see Figure 11). We then preprocess data by checking stationarity using ACF plots and PACF plots and found out that the average log-return data was pretty much stationary, then we fitted that data to our model and found out that our estimates were pretty much close to empirical estimates in **Analysing Dataset** section and also calculated standard errors and visualised Asymptotic regimes using estimated parameters (See Figure 12).

From our study we infer that Indigo and Spicejet log-return data is have moderate dependence in upper tail and lower tail is more towards independent side. Also Upper tail falls more in Asymptotic Dependent regime while Lower tail falls in Asymptotic Independent regime.
Practical perspective, our results could help to detect market risk contagion and be a useful source of information for investors who seek to diversify their portfolio.
Source Code Our Implementation can found at following github link :- Link

Numerical stability and further improvements

As We can see that the expression of joint distribution and density function of random Vector $\mathbf{X}$ consists integration which can not be solved to get explicit expression so numerical integration techniques were used and F_X contains exponential terms and its inverse can not be calculated explicitly for that various optimization techniques were used to numerically calculate inverse. So in these operations there are higher chances of getting numerical error or divergence in optimising techniques while implementing. Though we applied Logit transformations and did calculations on Log scale to minimise error but there can still be chances to improve these numerical calculations

and get more accurate results.
Although we focused in this paper on the bi variate setting, it can also be extended to multivariate conceptually but inference would hard.
Moreover, while we have here assumed that W has a Gaussian, it could be replaced by any other copula model that is asymptotically independent in both tails, without affecting the asymptotic tail results. Thus, the model construction is quite general and could be extended to a wide range of more complex and flexible copula models.

9 References

[1] Gong, Yan & Huser, Raphael. (2019). Asymmetric tail dependence modeling, with application to cryptocurrency market data.

[2] Paul G. Hoel, Paul Gerhard Hoel, Sidney C. Port, Charles J. Stone, Charles John Stone Houghton Mifflin, 1971

[3] R Documentation - https://www.rdocumentation.org/

[4] Investopedia.com - https://www.investopedia.com/terms/c/copula.asp

[5] https://en.wikipedia.org/wiki/Tail_dependence

[6] https://en.wikipedia.org/wiki/Copula_(probability_theory)

[7] https://youtu.be/oFUbYMFqSP8

[8] https://youtu.be/87iwrGs_KRA

[8] https://otexts.com/fpp2/stationarity.html