Everybody, welcome. Today, I will tell you about

what I think is a pretty useful application of copula models,

a multivariate method we'll be covering in week 6.

We'll be using them to model stock returns in order to build a portfolio and optimise it.

The data we'll be looking at, multivariate data,

are returns from five of these stocks: IBM,

Microsoft, British Petroleum, Coca Cola, and Duke Energy.

In practice, we will probably have more stocks,

but for our purposes, this will suffice.

These are the prices of these stocks over time in 2018 in particular.

As you can see, they go up.

Sometimes they go down rather quickly and they're correlated.

Now, when we run the returns,

we're mostly interested in return investment.

That's the amount we get divided by the amount we invest minus 1.

In particular, we can write down the daily return on investment.

That's the price of a stock today versus price yesterday minus 1.

That's how much money we would earn if we bought the stock yesterday,

sold it today, and ignoring transaction fees.

For example, if yesterday the stock was worth $100,

today 105, we would make five percent on our investment.

Now, it turns out that it's actually more

convenient to work with these data on the logarithmic scale.

That is, instead of looking at price today divided by yesterday minus 1, we'll look at the log,

usually natural logarithm, of price today divided by price yesterday.

Why would we want to do that? Well, consider.

We can write the return over two days,

that is log of price today over price two days ago,

as log of price today over price yesterday plus log of price yesterday

over price two days ago because the sum of logs is the log of a product.

What that lets us do,

is it lets us model log returns over long periods as some of the log returns over

short periods and it's much

more convenient to work with sums than with products as it turns out.

This is what our daily look like.

Notice that here we have several kinds of plots in one big plot.

We have the distributions of the stock.

Daily returns on the diagonal are the log returns.

On the diagonal, we have scatter plots showing

the pairwise relationships between the returns down here,

and then up here we have the correlations between the stocks.

You can see some of them are more correlated than others.

For example, IBM is heavily correlated with Microsoft.

On the other hand, Microsoft and Duke Energy are not correlated at all in practice.

Another issue is that we have

quite a few outliers in our data and sometimes long tails in particular,

sometimes stocks just drop by a lot.

That means that we cannot use normality in our models.

Instead, we'll use something called copula methods.

The basic idea is to use a tool called the inverse probability integral transform.

Now, that's a mouthful, but really what it does is it takes quantiles of the data,

it converts them to uniform quantiles and then it uses

the copula function to couple these uniform distributions together.

That is, we can consider the original distribution of the data,

normal or not, irrelevant.

We can make it irrelevant using this method.

We'll talk more about this in week 6.

But here's what it looks like once we've transformed

the data to uniform using this copula.

Notice that the distributions are on uniform of course,

but the correlations are preserved,

at least mostly preserved very close to what they were before.

Then we fit a copula model to this

and then we can simulate arbitrary realisations of these data.

That lets us experiment.

For example, with our simulated data,

we can try different portfolio strategies.

Symbols portfolio would perhaps be that for every dollar we invest,

we put 20 cents into each of the five stocks.

Now, it turns out that if we do that,

the average daily return will be 0.013 percent with standard deviation of 0.9.

I would expect to lose money on 47.5 percent of

the days and we'd expect to lose significant amount,

say more than one percent on about 11.5 percent of the days.

We can even plot the density of daily returns.

You can see it's largely centred at zero, but you know,

there is a slight bias upwards so we can make money off this.

But maybe we could do better than that.

We know that some stocks are correlated, so maybe we should hedge,

we should maybe not include redundant stocks.

In particular, can we get this return,

the 0.013 percent return we had the last time more reliably?

After some mathematics, which we'll take a look at in week 6,

we get these weights.

For every dollar we invest,

we invest 12.4 cents in IBM,

nothing in Microsoft, 21.4 cents in BP and so on.

Notice that because IBM and Microsoft were so highly correlated,

well, we just took one of them because it pretty much works as the other.

We didn't necessarily take the one that had the higher return because then

we'll be just chasing returns after they happen and we don't necessarily want to do that.

Average daily return from this is still 0.013.

But the standard deviation is now a little bit lower.

It's now 0.74 percentage points.

The probability of the loss is a little bit lower as well and the probability of a significant loss has shrunk quite substantially.

We'll look at this example in more detail and how to implement this in week 6.

In the meantime, welcome to multivariate analysis at UNSW.