Updated discussion 10/16/2017

0) By "distribution", do we mean "probability density function" or "cumulative distribution function"?

There is a famous theorem by Abe Sklar (Links to an external site.), that provides the theoretical basis for the study of copulas:
"Every multivariate cumulative distribution function, of a random vector $(X_1, \ldots, X_n)$ can be expressed in terms of its marginals distributions and a copula":

$$F_X(x_1, \ldots, x_n) = Pr(X_1 \leq x_1, \ldots, X_n \leq x_n) = C(F_1(x_1), \ldots, F_n(x_n))$$

where C is the copula and $F_1, \ldots, F_n$ are the marginals distributions of $X_1, \ldots, X_n$ respectively, that is $F_i(x) = Pr(X_i \leq x)$

In this context, distribution means cumulative distribution function.

1) Why do we need copulas if we can just compute correlations?

Certainly if you want to describe the joint distribution between random variables the information contained in the correlation matrix between these variables is not enough (it only contains information up to second moments). Thanks to a theorem due to Abe Sklar (see above) copulas are an essential ingredient to describe joint distributions, they can be thought of as what is left after marginal distributions are factored out.

2) What is the unique functional form of the univariate marginal distribution of a multivariate CDF that is a copula?

Since a copula is a multivariate CDF whose univariate marginal distributions are all Uniform(0,1), the functional form of the marginal distribution is the CDF of a uniformly distributed random variable on (0,1), that is F(x)=x. The density function is the constant 1 on (0,1) and 0 elsewhere.

Given any random variable X, you can define its cumulative distribution function as $F_X(x) = Pr(X \leq x)$, and for almost all random variables (nice random variables) this CDF is differentiable and its derivative $F'_X(x) = f_X(x)$ is called the probability density function of X.

There is a very well know transformation that allows to transform any random variable XX into a uniform random variable UU, as long as you know F_XFX (the cumulative distribution function of X X) and this happens to be a continuous function then, $U = F_X(X)$U=FX(X)is uniformly distributed on $(0, 1)$(0,1). The reason why UU is uniformly distributed is the following:

$$Pr(U \leq u) = Pr(F_X(X) = u) = Pr(X = F_X^{-1}(u)) = F_X(F_X^{-1}(u)) = u$$

Pr(U≤u)=Pr(FX(X)=u)=Pr(X=FX−1(u))=FX(FX−1(u))=u

Since $F_U(u) = u$FU(u)=u, this implies that $f_U(u) = 1$fU(u)=1 on $(0, 1)$(0,1) and 0 elsewhere, that is why the third plot on the sixth slide is the way it is.

This transformation is very useful in order to generate random variables having a particular CDF FF, out of a uniformly distributed random variable $X = F^{-1}(U)$X=F−1(U).

The probability of a single stock's return on any single day can be described by its probability density function, where there is typically a high probability around the mean and ever decreasing probabilities as you move toward the extreme high/low values in the tails -- looking something like a normal or t-distribution.

The cumulative distribution function shows the probability of any one day's return being less than a certain value. There is a 0% probability that it will be less than the lowest value in the left tail and a 100% probability that the value will be less than the highest value in the right tail. It increases from left to right.

The range of values from the CDF is always 0 to 1. Therefore, the probability that any one stock's CDF ranges from 0 to 1 is 100% for all stocks. Thus the PDF of the CDF is uniform -- always 1.

3) What does a correlation matrix corresponding to an independence copula look like? What about that corresponding to a co-monotonicity copula?

A correlation matrix corresponding to an independence copula would have 0's in all the non-diagonal cells. The correlation matrix corresponding to a co-monotonicity copula would have 1's in all of the cells.

4) Does a Gaussian copula imply that the variables have a multivariate Gaussian distribution? If not, what is that multivariate distribution called?

A Gaussian copula does not imply that the constituent variables have a multivariate Gaussian distribution nor that all of the univariate marginal distributions are Gaussian. A Gaussian copula indicates that the co-dependence between the random variables can be summarized by a correlation matrix - no higher moments are needed. If the multivariate data has a Gaussian copula it is called a meta-Gaussian distribution.

5) Suppose we have two stocks A and B with some bivariate distribution of returns. If stock A's return on day t is higher than its return on day s, then stock B's return on day t is also higher than its return on day s. Which correlation coefficient must be positive?

The Pearson correlation coefficient is $\langle (r_A(t) - r_A(s))(r_B(t) - r_B(s)) \rangle$, where $\langle \cdot \rangle$ represents an average over all t and s. Since we assume that the difference in returns in the two parentheses always have the same sign, this correlation coefficient must be positive.

If stock A's return on day t ranks high among its sample returns, then stock B's return on day t also ranks high among its sample returns. Which correlation coefficient must be positive?

Spearman rank correlation coefficient. The reasoning is same as the above for Pearson correlation: just interpret $r_A(t)$ to mean the rank of the return of stock A among its sample returns, and similarly for $r_B(t)$.

By the same reasoning, Kendall's Tau must also be positive also. Based on our assumption, we cannot have a discordant pair. Hence the probability of a discordant pair is zero. Since Kendall's tau is the probability of a concordant pair minus the probability of a discordant pair, it must be positive in this case.

6) Why doesn't it make sense to use Gaussian copula to model tail dependences as in "The Formula That Killed Wall Street?" (See http://archive.wired.com/techbiz/it/magazine/17-03/wp_quant?currentPage=all (Links to an external site.)). What might be a better copula for this measurement?

The Student's t-copula allows for joint fat tails and an increased probability of joint extreme events compared with the Gaussian copula.

Trent quoted from Wilmott: "Consider the share prices of two sneaker manufacturers: When the market for sneakers is growing, both companies do well and the correlation between them is high. But when one company gets a lot of celebrity endorsements and starts stealing market share from the other, the stock prices diverge and the correlation between them turns negative. And when the nation morphs into a land of flip-flop-wearing couch potatoes, both companies decline and the correlation becomes positive again."

Lucas shared his industry experience and insights gained during the financial crisis: "In the finance world, especially working for large publicly companies, technical and quantitative experts may not have power to be risk-averse or in this example, change Gaussian copula to t-Copula. For example, prior to financial crisis, normal distribution have been widely used in the finance world, such as Blacks-Sholes model. They are all based on the normal distributions. After financial crisis, we know that normal distribution assumption may not be a good indicator to withstand risk/heavy tails. I still remember I had a chance talking to several executives from IndyMac Bank in Pasadena, California before its failure. Several executives mentioned that they may foresee financial crisis but they cannot stop the bank to aggressively being subprime and Alt-A originators. If they stop doing that, or even slow down on that business, it would undermine their business strategy as well as their earnings, which will affect stock price. Shareholders, board and many other executives will be mad at that. Therefore, many finance people still want to generate high return and neglect the risk. Similarly to the example of using Gaussian copula, I think Mr. Li definitely know t-Copula is better for heavy tails. However, switching to t-Copula will cause lots of unnecessary pressure from the organization, executive leaders and the investors, who are only focus on stock appreciation. The reason is that assumption change will affect company's earnings, such as set up more provisions each year and this will adversely affect net income or earnings, which will ultimately drag down the stock price."

Jonah commented that “Apart from lack of sufficient tail dependence, the article also emphasizes two additional shortcomings of Li’s model. First, Li’s model was parameterized using prices of Credit Default Swaps rather than historical correlation data. While the approach was clever, he had a limited history of CDS prices (~10 years) and the period of prices he did have represented generally stable/favorable economic conditions. Secondly, Li made the assumption that the complex correlation relationship could be captured by a single value, rather that a dynamic parameter that is likely necessary.”

7) Correlations only capture the second moment of distributions, so we need copulas. But to be honest, I'm still trying to understand what moments are. Why is it important to capture the second moment?

Ans. To find any dependence between 2 random variables, we have to multiply them together. Whenever we multiply 2 random variables together, we get the second moment. Hence second moment is important, and it is called "covariance".

The first moment is equal to the mean (it measures "central tendency"). The second central moment is equal to the population variance (it measures "dispersion"). The third moment measures how "lopsided" the data is (it is equal to skewness after standardizing the data.) The fourth moment measures how heavy the tails are (kurtosis.) And you could have even higher moments in order to capture more complex attributes of the data.

Having all the moments helps you to determine (or to model) the correct distribution of your data. For instance, if your data is normally distributed, then you only need the first and second moment. But if your data is skewed or has heavy tails (like the t distribution) then you will need higher moments too.

If I try to fit data that is heavily skewed (for instance house prices or income) to a "normal distribution", I will certainly get a model (e.g. I will get values for my mean and variance parameters) but this model is going to be a very bad one and could lead me to errors. If accuracy is important, I need to fit a model that supports higher order moments (skewness, etc.) in this case.

That was for a univariate distribution. For a bivariate distribution moments exists too. The second moment would be the covariance or correlation which measures how much the variables "linearly agree". And you can have higher moments for instance to measure how their extreme values interact (e.g. the behavior in the tails of the distribution), etc.

Like in the univariate example, if I use a copula such as the Gaussian that only has one parameter (the correlation matrix) but my variables have different interactions in the center vs. in the tails, non-linear relationships, or other complex interactions, then my Gaussian copula will be representing the data very poorly. In that case I need a copula that captures higher moments (with more parameters) in order to have an accurate model. To prof Chan's question, I should not ignore those higher moments.

8) Ashwini asked “could you please give us some real world examples where estimates of Copulas and Tail Dependencies can be used? How the decisions can be made using those estimations? How do we use the estimated copula parameter like rho [the Spearman’s rank correlation] to make informed decision and what is the use of tail dependency coefficient?”

Ans. If you are managing a portfolio and you find that the tail dependence index is high, then one should not invest too much in that portfolio. When there is a big market movement, the value of the portfolio may drop a lot because all the stocks will drop together. Diversification won’t help in this case. Of course, how much is too much depends on your risk tolerance and the Value-At-Risk of the portfolio, which will be covered in a later module. So tail index can be thought of as a warning bell. You can, of course, arbitrarily decide not to invest in a portfolio with a tail index exceeding some threshold.

Rank correlation is often a more robust measure of co-dependence than the usual Pearson correlation coefficient. So you can use that as an indicator for many trading strategies which require correlation as an input. For example, in the last chapter we will be discussing statistical factors, which depends on the covariances of returns as an input. But if you know that the returns have a high degree of tail behavior, you may substitute covariances with rho times each stock's individual variance.

9) Vikrama asked “I am also finding difficulty how copulas will be helpful in decision making.”

Ans. Just finding the correlation, covariance, or copula are not the final goal of copula models. Once has to incorporate these values in a Value-At-Risk or Estimated Shortfall risk assessment, both to be discussed in Module 7. Also, as I mentioned to Robert, these correlations or copulas can also be used as input to a portfolio optimization program to find the optimal asset allocation in order to maximize returns while minimizing risks. This process will be discussed in Module 8.

So you see, one thing lead to another in this course: they are all interrelated!

10) Robert asked “how much more difficult is the modeling and math associated with this week’s copula topic for multivariate models with more than 2 variables (i.e say a portfolio of 100 stocks)?... I still have a similar concern that I raised last week: that correlations and relationships between variables change over time. So it's not clear to me how one relies on a model like a copula to consider variable dependence when such dependence is dynamic. Perhaps put differently, I'd be curious what model risk management system can be put in place to register a warning signal to the trader that variable co-relationships are breaking down and thus perhaps the model in production needs to be recalibrated.”

Ans. It is no more difficult to find a copula model for 100 or 1,000 stocks than a copula for 2 stocks. The only constraint is your computer's CPU and memory size! For e.g. if you look at https://www.rdocumentation.org/packages/copula/versions/0.999-18/topics/fitCopula. you will see that it can handle an input data matrix with arbitrary dimension $n \times d$.

The current applications of copulas are still mainly in risk management and portfolio optimization. In portfolio optimization, we need to find the best trade-off between risk and returns. One needs copula models to better determine both risks and aggregate returns due to the co-dependence of stock returns. This is a topic for Module 8.

But it is true that most models are backward looking. Even ostensibly dynamic models are backward looking, since all parameters of such models need to be optimized with historical data. Even human discretionary traders have to use their past experience to inform their "intuition" or "judgement" on the possible future market condition. Such judgment may be no better than what a comprehensive algorithmic trading model can predict, if by "comprehensive" we meant it incorporates all possible financial, economic, and political variables that a human can use, and use techniques of "deep-learning" to synthesize such variables.

Copula models do not pretend they can signal impending danger. They are created to reflect more accurately what the current market risks are, not to predict what future market risks will be. They are descriptive, not predictive.

10) Robert asked: “In code line 23 on page 211 in Ruppert (the code provided to answer Problem 6), one of the arguments is ft1@estimate. I've not before come across that @... syntax in R. Does anyone have any thoughts on what exactly that does / when it is best used?”

Ans. Here are some references:

http://astrostatistics.psu.edu/su07/R/html/base/html/slotOp.html (Links to an external site.)Links to an external site.

or

http://astrostatistics.psu.edu/su07/R/html/methods/html/slot.html (Links to an external site.)Links to an external site.

Basically, Object@name extracts the field called "name" from an Object.

11) Scott asked “In the several asset managers I've worked for and the numerous I've performed due diligence on, CAPM is not really discussed much. The firm's I've worked for have been primarily fundamental though, where the investment approach is somewhat unsystematic (at best). How would you describe your use of CAPM in practice?”

My answer: “CAPM is a first order approximation of what the returns of an asset should depend on (i.e. proportional to the market index return, with "beta" as the proportionality constant.) Naturally, this is too simplistic a model! So the problems are too numerous to discuss, since it is a very crude and unrealistic model. In the last module, we will discuss factor models which is second order approximation, and a bit more realistic.

The way people use CAPM is to focus on the beta of an asset. If they have to characterize an asset's returns and risks characteristics with one number, beta will be that number. A high-beta stock means it is high return and high risk. Beta is

Corr(asset_return, index_return)*sd(asset_return)/sd(index_return),

where Corr is correlation, sd is the standard deviation. So sd(asset_return) is the volatility of the asset. You can verify that this expression is same as Ruppert Eq. 17.7, but I find this version more intuitive. It tells us that beta is partly correlation to the market, but partly how volatile the asset is compared to the market.

Finance practitioners almost never talk about CAPM - that's for professors and armchair academics. We barely can remember what CAPM means (I had to look up the definition myself every time someone asks!) But we know in our sleep what beta is. If we own a high beta stock, we won't sleep well.