Article can be accessed online at http://www.publishingindia.com

A Study of the Copula Approach to Pair Trading

Apoorva Uday Nayak*, Megha Ugalmugale, Vishwanjali Gaikwad***, Vahida Z. Attar******

Abstract

The stock market is one of the most important sources for companies to raise money. The prediction of stock prices has always been a challenging task. Investors are always in search of profitable trading strategies which provide accurate trading signals. Pair trading is an investment strategy of matching a long position (buy) with a short position (sell) in two historically correlated stocks. In pair trading, the use of correlation or co-integration as a dependence measure is central. Some traditional techniques also assume symmetric distribution of data along the mean zero. However, in case of financial assets the occurrence of symmetric distributions is very rare and, hence, the use of these techniques may lead to erroneous results. Copula is a relatively new pair trading technique which describes dependence structures for linear and non-linear distributions without making rigid assumptions. Copula, thus, overcomes the shortcomings of traditional pair trading techniques and gives accurate trading signals. This paper is a study of different types of copulas and their application in finance to fit different types of data.

Keywords: Copula, Pair Trading, Stock, Linear Dependence, Normal Distribution, Kendall Tau

1. Introduction

Pair trading is an investment strategy. It was developed to generate significant and consistent returns while controlling risk. It is a market neutral strategy meant to generate profit regardless of whether equities rise or fall. The strategy of matching a long position (buy) with a short position (sell) in two historically correlated stocks is known as pair trading. The buying of a security such as a stock with the expectation that the asset will rise in value is known as long position whereas the sale of a borrowed security with the expectation that the asset will fall in value is known as short position.

The temporary weakening of correlation between the two paired stocks is the point of interest in pair trading. This can be caused by temporary changes in demand and supply or large buy and sell orders for one of the paired stocks. The divergence between paired stocks can also be caused due to one company being in the news for some important reason. Pair trading takes advantage of this temporary mispricing of the assets and hence, categorized as statistical arbitrage trading strategy. In this scenario, one stock is observed to go up while the other stock is observed to fall. Pair trading involves going short on the outperforming stock and going long on the underperforming one with the expectation that the paired stocks will converge over time. Hence, pair trading is a convergence trading strategy (Stander, Marais & Botha, 2013). For example, consider two companies A and B belonging to the same sector. Ideally, both A and B should observe similar rise or fall depending on their demand in the market. However, if the price of stock A were to go up significantly while B observed no change, the pair trading strategy would suggest buying stock B and selling stock A. If the price of stock B rose to close that gap in price, the trader would make money on stock B, while if the price stock A fell, he would make money on having shorted the stock A. Pair trading is thus a mean-reverting strategy with the expectation that prices will eventually

* Student, College of Engineering, Pune, Maharashtra, India. E-mail: apoorvanayak29@gmail.com
** Student, College of Engineering, Pune, Maharashtra, India. E-mail: meghau1992@gmail.com
*** Student, College of Engineering, Pune, Maharashtra, India. E-mail: vishwanjali12692@gmail.com
**** Assistant Professor, Department of Computer Engineering, IT at the College of Engineerig, Pune, Maharashtra, India. E-mail: vahida.comp@coep.ac.in

return to their historical trends. It is a market-neutral strategy which helps reduce the market risk. Even if the market plummets on a particular day and the two stocks move down with it, the trade results in a gain on the short position and a loss on the long position which is negated thus leaving the profit close to zero in spite of the large move.

Overview

Copula is a statistical measure that represents a multivariate uniform distribution. It examines the association or dependence between random variables. Copula is a mathematical tool that can be used to solve many financial problems. However this concept wasn't used in finance till the late 1990s. The notion of copula was first introduced in Sklar's Theorem which states that any multivariate joint distribution can be written in terms of univariate marginal distribution functions and a copula which describes the dependence structure between the variables (Salinas-Gutirrez et al., 2010).

For example, let H (x, y) be a joint distribution of two random variables X and Y with marginal distribution functions F and G. Then there exists a copula C such that

$$H(x, y) = C(F(x), G(y))$$

The converse also holds true. For any univariate distribution functions F and G and any copula C, the function H is a two-dimensional distribution function with marginal F and G. If F and G are continuous, then C is unique. Practical advantage of copula-based approach to modelling is that appropriate marginal distributions for the components of a multivariate system can be selected freely, and then linked through a suitable copula. That is, copula functions allow one to model the dependence structure independently of the marginal distributions. In pairs trading, the assumption of and the use of correlation or co-integration as a dependence measure are central. Some traditional techniques also assume symmetric distribution of data along the mean zero. However, the occurrence of non-linear distributions is quite frequent in financial assets. Negative skewness and/or excess kurtosis are also observed. This results in lower and upper tail dependencies. Outside the world of elliptical distributions, the use of linear correlation coefficient as a dependence measure is erroneous and may lead to misleading results. This may trigger wrong trading signals and may fail to recognise profit opportunities. Copulas are much more realistic. A significant shortcoming which they overcome is the non-assumption of normal distribution for all types of data and the use of different dependence measures other than the linear correlation coefficient. Copula can be applied regardless of the form of marginal distribution thus providing much more robustness and flexibility in practical applications. Modelling asset returns is one of the most important problems in finance. Copula provides a powerful framework for describing dependence structures without rigid assumptions. Copula measures lower and upper tail dependencies, considers linear and non-linear relationship and results in far richer set of information with the shape and nature of dependency between the stock pairs. Copula is invariant under strictly monotonic transformations and hence the same copula is obtained regardless of the form of data i.e. price series and price series when converted to return series have the same copula. The two main steps involved in the pair trading with copula technique is the selection of an appropriate copula for the stock pairs and identification of trading opportunities and relative positions between stock pairs. This separation of procedure is important as it provides greater flexibility and the analyst can use different marginal distributions to resolve the diversity in financial risks (Ane & Kharoubi, 2003). Decomposing the joint distribution into individual marginal distributions and a copula helps as it allows for the construction of better models of the individual variables than would be possible if only explicit multivariate distributions were considered. Applying the best-fitting marginal distribution ensures that all information regarding the dependence structure between random variables are accurately captured before estimating its joint distribution without rigid assumptions (Aas, 2004; Liew & Wu, 2013).

Types of Copula

Consider the two well-known parametric families of copula: Elliptical copula, Archimedean copula

Elliptical Copula

Elliptical copulas are those relating to elliptical distributions. They are also known as implicit copulas. These copulas follow linear dependence structure. The linear correlation coefficient is used to measure their linear dependence. They do not have simple closed forms but are extracted from multivariate distribution functions using Sklar's theorem. The asset returns which are well represented as elliptical distributions belong to elliptical copulas. Most common elliptical copulas are Gaussian and Student-t copulas.

(1) Normal Copula: Normal copula or the Gaussian copula is the copula of the multivariate normal distribution. The distribution of a normal random variable with mean 0 and variance 1 is called standard normal distribution. It is a bell shaped curve. Normal copula Cp(u, v) is given as

$$\int_{-\infty}^{\phi-1}(u)\int_{-\infty}^{\phi-1}(v)\frac{1}{2\pi(1-\rho^2)^{1/2}}\exp\left\{\frac{x^2-2xy+y^2}{2(1-\rho^2)}\right\}dx\,dy$$

The formula for $P(U \le u|V = v)$ is given as

$$\Phi\left(\frac{\Phi^{-1}(u)-\theta\Phi^{-1}(v)}{\sqrt{1-\theta^2}}\right)$$

The formula for $P(V \le v|U = u)$ is given as

$$\Phi\left(\frac{\Phi^{-1}(v)-\theta\Phi^{-1}(u)}{\sqrt{1-\theta^2}}\right)$$

Here,

ϕ is cumulative distribution function and θ is the correlation coefficient given as

$$r_{xy}=\frac{n\Sigma x_i y_i-\Sigma x_i\Sigma y_i}{\sqrt{n\Sigma x_i^2-(\Sigma x_i)^2}\sqrt{n\Sigma y_i^2-(\Sigma y_i)^2}}$$

$\theta = 1$ indicates perfect positive correlation

$\theta = 0$ indicates no correlation

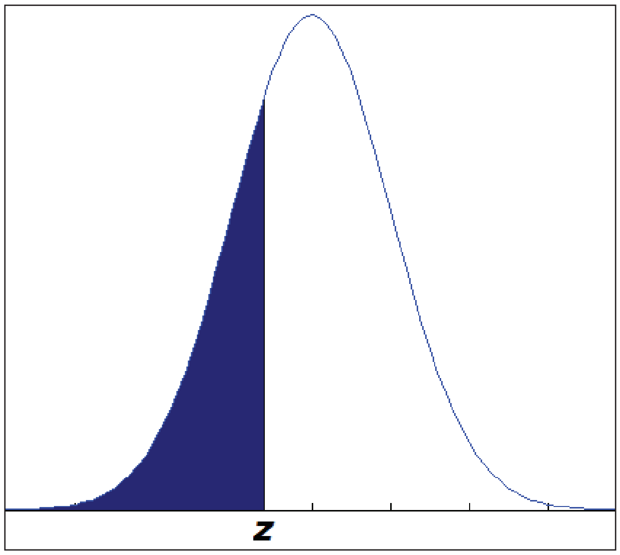
$\theta = -1$ indicates perfect negative correlation

For a random variable X cumulative distribution function is given by $F(x) = P(X \le x)$ which is the probability that X takes a value less than or equal to x. For continuous distributions, like in the case of normal distribution, it gives the area under the probability density function from $-\infty$ to x. All observations of any random variable X can be converted to observations of a normal random variable Z with mean 0 and variance 1. This can be done by

$z = (x - \mu)/\sigma$

When X takes the value x, Z will take the value z given by the above formula. This z when looked up in the normal probability table gives the area under the curve to the left of z (Fig. 1) which is the probability of finding random variable X less than or equal to x (Walpole et al., 2005).

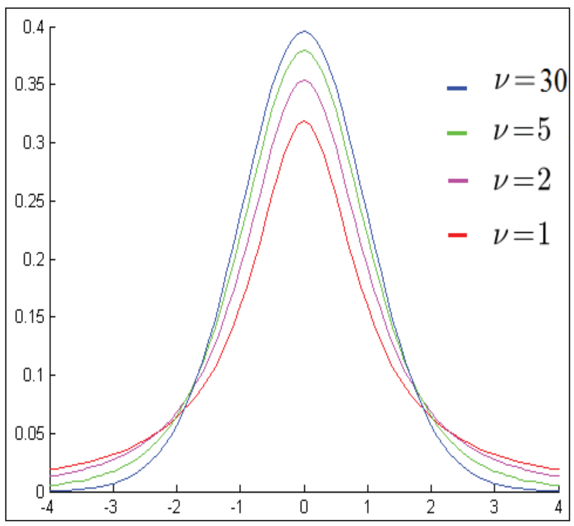
Fig.1: Normal Distribution with Shaded Area Representing P(X <= x)



$\Phi^{-1}(u)$ is the inverse of the standard univariate Gaussian distribution function. It gives the z value corresponding to which lies an area of u to the left in the normal probability curve.

(2) T Copula: T distribution is a type of elliptical distribution. It is a bell shaped curve and is similar to normal distribution. However, it is more variable as T values depend on fluctuations of two quantities x (mean) and s (Sample Variance) whereas the z values depend only on the changes of x from sample to sample. The variance of T depends on the sample size n and is always greater than 1. For $n \ge 30$ the distribution of T does not differ much from the standard normal distribution and the T copula also converges to normal copula (Fig. 2).

Fig.2: Student-T Distribution with V Degrees of Freedom



Formula for T copula Cp,v(u, v) is given as

$$\int_{-\infty}^{t_v^{-1}}(u)\int_{-\infty}^{t_v^{-1}}(v)\,\frac{1}{2\pi(1-\rho^2)^{1/2}}\left\{1+\frac{x^2-2xy+y^2}{2(1-\rho^2)}\right\}^{-(v+2)/2} dx\,dy$$

In contrast with the Normal copula, Student-T copula can model tail dependence. It allows for joint fat tails and an increased probability of joint extreme events compared with the Gaussian copula. Two parameters namely linear correlation and degree of freedom are used to capture the correlation between the risk classes. Increasing the value of degree of freedom decreases the tendency to exhibit extreme co-movements. Hence, Student-T copula is generally considered a better solution to model the dependence structure for operational risk data.

The formula for $P(U \le u|V = v)$ is given as

$$t_{\theta 1}+1\left(\sqrt{\frac{\theta_1+1}{\theta_1+t_{\theta 1}-1(v)^2}} * \frac{t_{\theta 1}^{-1}-\theta_2 t_{\theta 1}-1(v)}{\sqrt{1-\theta_2^2}}\right)$$

The formula for $P(V \le v|U = u)$ is given as

$$t_{\theta 1}+1\left(\sqrt{\frac{\theta_1+1}{\theta_1+t_{\theta 1}-1(u)^2}} * \frac{t_{\theta 1}^{-1}(v)-\theta_2 t_{\theta 1}-1(u)}{\sqrt{1-\theta_2^2}}\right)$$

Here,

θ_1 is the degree of freedom which is given as sample size-1

θ_2 is the linear correlation coefficient

$t_{\theta 1}(T)$ - gives the area α under the T distribution curve to the left of the T value for θ_1 degrees of freedom

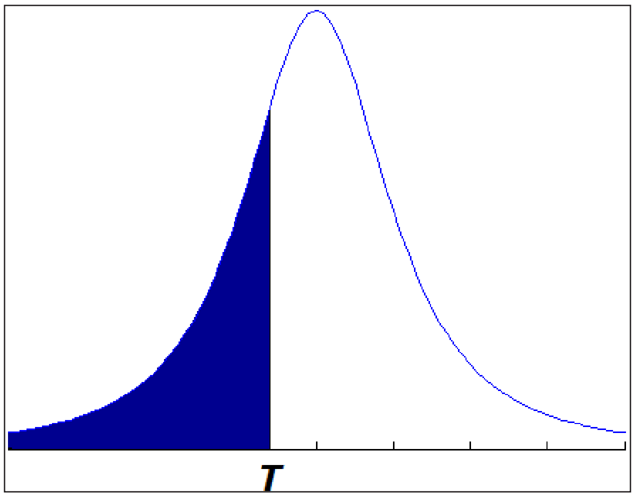
$T = \bar{x} - \mu/(S/\sqrt{n})$

$t_{\theta 1}^{-1}(u)$ is inverse of the standard univariate Student-T distribution with θ_1 degrees of freedom, expectation 0 and variance $\theta_1 / \theta 1 - {}^{2}$

It gives the T value that leaves area α to the left, as shown in Fig. 3, where $\alpha = t_{\theta 1}(u)$

The standard bivariate Student-T distribution has a disadvantage that both the marginal distributions must have the same tail heaviness which may not be true for stock returns. It does not account for asymmetries in data. Hence, we need to study another class of parametric copulas which is the Archimedean copulas.

Fig.3: T Value that Leaves Area A to the Left



Archimedean Copula

One of the key disadvantages of elliptical copulas is that they do not have closed form expressions. Archimedean copulas, on the other hand, are not derived from multivariate distribution functions but admit an explicit formula. They are also known as explicit copulas. Archimedean copulas have non-elliptical distribution and follow non-linear dependence structure. They describe stronger dependence between extreme data. They can model dependence with only one parameter monitoring the strength of dependence even in high dimensional cases. They make use of Kendall's Tau for calculating dependency. Although restrictive for higher dimensional cases, they fit bivariate return distributions extremely well. As the project demands analysing stocks in pairs, only bivariate distributions are generated and the use of Archimedean copulas is completely justified. Most common Archimedean copulas are the Clayton, Gumbel and Frank copulas.

Kendall's Tau

Kendall's Tau coefficient is a statistic used to measure of the strength of dependence between two quantities. It reflects the association between the two quantities when ranked by each of the quantities. Consider two samples X and Y each of size n and each containing distinct elements.

Let $X = \{x_1, x_2, x_3, ...,x_n\}$ and $Y = \{y_1, y_2, y_3, ..., y_n\}$ $(x_1, y_1),(x_2, y_2),(x_3, y_3), ...,(x_n, y_n)$ are the joint set of observations for X and Y .

The total number of possible pairings between the observations of X and Y is $n^*(n-1)/2$

The Kendall's Tau coefficient is defined as

$$T = \frac{(\text{number of concordant pairs}) - (\text{number of discordact pairs})}{\frac{1}{2}n(n-1)}.$$

where,

a pair $(x_i, y_i),(x_j, y_j)$ is concordant if $(x_i > x_j$ and $y_i > y_j)$ or $(x_i < x_j$ and $y_i < y_j)$;

a pair (x_i, y_i), (x_j, y_j) is discordant if $(x_i > x_j$ and $y_i < y_j)$ or $(x_i < x_j$ and $y_i > y_j)$

As the denominator represents the total number of possible pairs, the value of Kendall's Tau should lie in between $-1 \leq \tau \geq 1$.

Properties of Kendall's Tau:

(1) Kendall's Tau value of 1 indicates positive association between the two quantities

(2) Kendall's Tau value of -1 indicates negative association i.e. one ranking is the reverse of the other

(3) Kendall's Tau value of 0 indicates no association i.e. the two quantities are independent of each other

For example

S. No.	X	Y
1	10	11
2	12	09
3	18	15
4	20	04
5	11	19

Consider the pair (x_1, y_1) and (x_2, y_2)

$(x_1, y_1) = (10, 11)$

$(x_2, y_2) = (12, 09)$

As $10 < 12$ and $11 > 09$, (x_1, y_1) and (x_2, y_2) forms a discordant pair.

Now consider the pair (x_1, y_1) and (x_3, y_3)

$(x_1, y_1) = (10, 11)$

$(x_3, y_3) = (18, 15)$

As $10 < 18$ and $11 < 15$, (x_1, y_1) and (x_3, y_3) forms a concordant pair.

Continuing in this way, 10 pairs can be formed. In this example, the total number of concordant pairs is 3 while the number of disconcordant pairs is 7.

Therefore, Kendall's Tau is 3−7/10 = −0.4. We have considered only distinct values in the above example. However, if the elements of X or Y are not distinct, Tau-b is used to calculate association. Consider that the fourth element of X is 18. Here, for the pair (x_3, y_3) and (x_4, y_4), $x_3 = x_4 = 18$.

Hence, the above pair is neither concordant nor discordant. For such cases Kendall's Tau is defined as

$$\tau_B = \frac{n_c - n_d}{\sqrt{(n_0 - n_1)(n_0 - n_2)}}$$

where,

n_c = Number of concordant pairs

n_d = Number of discordant pairs

$n_0 = n * (n-1)/2$

$n_1 = t_x * (t_x - 1)/2$

where, t_x is the number of tied elements in x

$n_2 = t_y * (t_y - 1)/2$

where, t_y is the number of tied elements in y. In this example, $t_x = 2$ (as 2 of the 5 values are same) $t_y = 0$ (as no same values for y) $n_1 = (2 * 1)/2 = 1$ $n_2 = 0$ denominator $= \sqrt{(10 - 1)(10 - 0)} = \sqrt{(9 \text{ r } 10)} = 9.48 = \text{p}$ τB = (3 − 6)/9.48 = −0.3162. Kendall's Tau measures the degree of monotonic dependence between the two samples and is invariant under monotone transformations. Linear correlation, on the other hand, measures only the degree of linear dependence and is generally not invariant under monotone transformations.

(1) Clayton Copula: Clayton copula is an asymmetric Archimedean copula. It exhibits greater dependence in the negative tail than in the positive and is defined as

$$C_\theta(u,v) = (\max\{u^{-\theta} + v^{-\theta} - 1; 0\})^{-1/\theta}$$

The formula for $P(U \leq u|V = v)$ is given as

$$v^{-(\theta+1)}\left(u^{-\theta} + v^{-\theta} - 1\right)^{-\frac{1}{\theta}-1}$$

The formula for $P(V \leq v|U = u)$ is given as

$$u^{-(\theta+1)}\left(u^{-\theta} + v^{-\theta} - 1\right)^{-\frac{1}{\theta}-1}$$

Here, the parameter θ defines the degree of dependence between the variables using Kendall's Tau and can be given as $\theta = 2 * \tau/1 - \tau$

$\theta \in [-1, +\infty)\backslash\{0\}$

$\theta = \infty$ indicates perfect dependence (Vose Software TM)

Clayton copula produces a tight correlation at the low end of each variable. It interpolates between perfect negative and perfect positive dependence (McNeil, Frey, & Embrechts, 2005).

(2) Gumbel Copula:Gumbel distribution is used to model the distribution of maximum or minimum of a number of samples of various distributions. If the maximum close price of each year of a stock is given then the gumbel distribution represents the distribution of the maximum closing price of the stock in a particular year.

Gumbel copula is an asymmetric Archimedean copula. It can model tail dependence. It exhibits greater dependence in the positive tail than in the negative and is defined as

$$C_\theta(u,v) = \exp(-((-\log(u))^\theta + (-\log(v))^\theta)^{1/\theta})$$

The formula for $P(U \leq u|V = v)$ is given as

$$C_\theta(u,v) * [(-lnu)^\theta + (-lnv)^\theta]^{\left(\frac{1-\theta}{0}\right)} * (-lnv)^{\theta-1} * \frac{1}{v}$$

The formula for $P(V \leq v|U = u)$ is given as

$$C_\theta(u,v) * [(-lnu)^\theta + (-lnv)^\theta]^{\left(\frac{1-\theta}{0}\right)} * (-lnv)^{\theta-1} * \frac{1}{u}$$

Here, the parameter θ defines the degree of dependence between the variables using Kendall's Tau and can be given as

$\theta = 1/1 - \tau$

$\theta \in [1, +\infty)$

$\theta = \infty$ indicates perfect dependence

$\theta = 1$ indicates independence

It produces more correlation at the two extremes of the correlated distribution but has its highest correlation in the maxima tails (Soprano et al., 2010).

(3) Frank Copula: Frank copula is an asymmetric Archimedean copula. It is defined as

$$C_\theta(u,v) = \frac{1}{\theta}\left(1 + \frac{(\exp(-\theta u)-1)(\exp(\theta v)-1)}{\exp(-\theta)-1}\right)$$

The formula for $P(U \leq u|V = v)$ is given as

$$\frac{e^{-\theta v}\left(e^{-\theta u} - 1\right)}{(e^{-\theta} - 1) + (e^{-\theta u} - 1)(e^{-\theta v} - 1)}$$

The formula for $P(V \leq v|U = u)$ is given as

$$\frac{e^{-\theta u}\left(e^{-\theta v} - 1\right)}{(e^{-\theta} - 1) + (e^{-\theta u} - 1)(e^{-\theta v} - 1)}$$

Parameter θ is given by $[D1\ (\theta)-1] / \theta = 1-\tau / 4$

$$\frac{[D_1(\alpha)-1]}{\alpha} = \frac{1-\tau}{4}$$

$D1(\theta)$ is Debye function with n = 1.

Debye function is given as

$$D_n(x) = \frac{n}{x^n}\int_0^x \frac{t^n}{e^t - 1}\,dt.$$

$\theta \in [-\infty, \infty) \setminus \{0\}$

Each of the above mentioned copula have their own set of advantages and disadvantages. Hence, choice of copula forms an important step. Selection of the best copula can be made by maximizing the likelihood function value.

Information Criteria

As mentioned earlier, the two main steps involved in the pair trading with copula technique are selection of an appropriate copula for the stock pairs and identification of trading opportunities and relative positions between stock pairs. Visually, the type of distribution (normal, student-t, clayton, gumbel, frank) into which the data fits can be determined by plotting the graphs and selecting the corresponding copula of the distribution which resembles the graph. However doing this computationally is not possible and some other methods other than visual aids are required.

The Chi-Squared, Kolmogorov-Smirnoff and Anderson-Darling are some of the goodness of fit statistics. They describe how well a statistical model fits a set of observations.

Although very popular, the above models are restricted to data with precise observations and are not useful for censored or truncated data. Hence, they are all inappropriate as a method of comparing fits of distributions to data. For important work, information criteria are used.

Information Criterion is a test which helps in selection of a model which that adequately fits data. We look at the following three commonly used information criterion.

(1) Akaike Information Criterion (AIC): It gives a measure of the relative quality of a statistical model for a given set of data. It offers an estimate of the information lost when a given model is used to represent data.

AIC = 2 * k −2ln[Lmax]

where,

k = number of parameters to be estimated (e.g. the Normal distribution has 1: linear correlation, Student-t distribution has 2: linear correlation and degree of freedom)

Lmax = the maximized value of the Likelihood (fit the parameters by MLE and record the natural log of the Likelihood) for the estimated model.

In practice, the AIC values of all candidate models are found using the above formula. The model with the lowest AIC value is the model with lowest information loss and is selected as the model which fits the data to the nearest. As k is the number of parameters, an increase in k increases the value of AIC. Hence, AIC not only values goodness of fit but discourages overfitting i.e. having too many parameters relative to the size of data.

(2) Bayesian Information Criterion (BIC): BIC is a criterion for model selection among a finite set of models. It is also known as Schwarz information criterion(SIC).

SIC = ln[n]k − 2ln[Lmax]

where,

n = number of observations (e.g. data values, frequencies)

k = number of parameters to be estimated (e.g. the

Normal distribution has 1: linear correlation, Student-t distribution has 2: linear correlation and degree of freedom)

Lmax = the maximized value of the Likelihood (fit the parameters by MLE and record the natural log of the Likelihood) for the estimated model.

Similar to AIC, the BIC values of all candidate models are found using the above formula. The model with the lowest BIC value is selected as the model which fits the data to the nearest. An increase in the number of parameters increases the value of likelihood (second term) but results in overfitting. Like AIC, BIC also penalizes overfitting, albeit more strongly with the first term of the expression. Thus, BIC discourages increase in complexity where complexity refers to number of parameters in a model.

(3) Hannan Quinn Information Criterion (HQIC): HQIC is also a criterion for model selection. It is an alternative to AIC and BIC.

HQIC = 2ln[ln[n]]k − 2ln[Lmax]

where,

n = number of observations (e.g. data values, frequencies)

k = number of parameters to be estimated (e.g. the

Normal distribution has 1: linear correlation, Student-t distribution has 2: linear correlation and degree of freedom)

Lmax = the maximized value of the Likelihood (fit the parameters by MLE and record the natural log of the Likelihood) for the estimated model. Similar to the above two information criteria, the HQIC values of all candidate models are found using the above formula. The model with the lowest HQIC value is selected as the model which fits the data to the nearest. Any one of the above criteria can be used to find the model which adequately fits the data. It can be observed that the above three formulae are closely related and all three of them employ the usage of maximum log likelihood.

D. Maximum Likelihood Estimation:

As the name suggests, the method of maximum likelihood is that for which the likelihood function is maximized. It provides an estimate to the parameters of the model. Likelihood is hypothetical probability in which the event which has already occurred is considered to yield a particular outcome. Probability and likelihood although used interchangeably are somewhat different. Probability refers to the occurrence of future events whereas likelihood refers to past events with known outcomes. For example, if a fair coin is tossed three times then probability is used to find if we get a head each time whereas if the coin was flipped three times and we got a head each time then likelihood is used to tell if the coin is fair. The probability of getting certain outcomes x given the parameter θ is the likelihood of that parameter θ given the outcomes x.

$P(x|\theta) = L(\theta|x)$

Given independent observations $x_1, x_2, x_3, ...,x_n$ of probability density/mass function $f(x; \theta)$ then the maximum likelihood estimator is the one which maximizes the likelihood function

$$L(x_1, x_2, x_3, ...,x_n; \theta)$$

$= f(x_1, \theta) * f(x_2, \theta) * f(x_3, \theta) * .....*f(x_n, \theta)$

$= P(X_1 = x_1, X_2 = x_2, X_3 = x_3, ...,X_n = x_n|\theta)$

which is the probability of obtaining the sample values $x_1, x_2, x_3, ..., x_n$.

For example, consider that three items are inspected to check whether they are defective or non-defective. Testing them results in two non-defective items followed by a defective item. To find the proportion p of non-defective items, the likelihood is given by

$p r p r (1 - p) = p^2 - p^3$

Maximum likelihood estimation would give an estimate of p for which the likelihood is maximized. If we differentiate the above expression and set the derivative to zero, we get

$p = 2/3$

So MLE is the reasonable estimator of a parameter based on sample information which is the parameter value that produces the largest probability of obtaining the sample. Many times it is convenient to work with the natural log of the likelihood function for finding the maximum of the likelihood function. The only change in this approach is to find natural logarithm of likelihood function before finding its derivative and setting it to zero.

Conclusion

Pair trading can seek profits from the difference in price change between two related instruments when the pairs are appropriately formed. The copula function instead of assuming normality of financial data, describes the dependency between the two instruments by relating the joint distribution with their most appropriate marginal distributions. Thus use of copula for implementing pairs trading exhibits high profits.

Acknowledgement

The authors wish to express their heartfelt thanks to Dr. Aniruddha Pant and Mr. Rohit Walimbe of Algo Analytics Financial Consultancy Pvt. Ltd. for their immense contribution in improving the quality of this paper.

References

1. Aas, K. (2004). Modelling the dependence structure of financial assets: A survey of four copulas. A Technical Report, NorskRegnesentral.
2. Liew, R. Q. & Wu, Y. (2013). *Pairs Trading*: *A Copula Approach*. In Proceedings of Asian Finance Association Conference.
3. McNeil, A. J., Frey, R. & Embrechts, P. (2005). *Quantitative risk management: Concepts, techniques, and tools*. Princeton University Press.
4. Salinas-Gutirrez, R., Hernndez-Aguirre, A., Rivera-Meraz, M. J. J. & Villa-Diharce, E. R. (2010).*Using Gaussian Copulas in Supervised Probabilistic Classification*. Proceedings of the 9th Mexican International Conference on Artificial Intelligence Conference on Advances in Soft Computing: Part II (pp. 104-115).
5. Soprano, A., Crielaard, B., Piacenza, F. & Ruspantini, D. (2010). *Measuring operational and reputational risk: a practitioner's approach*. Wiley Publications.
6. Stander, Y., Marais, D. & Botha, I, (2013). Trading strategies with Copulas. *Journal of Economic and Financial Sciences*, 6(1), 83-107.
7. Walpole, R. E., Myers, R. H. & Myers, S. L. (2005). Keying Probability and Statistics for Engineers and Scientists.