CHAPTER 2

Methodologies

All the methods are discussed in this chapter, such as CAPM with belief function, CAPM with interval data, and CAPM-based vine copulas. First, this chapter introduces the CAPM model with Belief functions. Second, we present Interval value in linear regression .Third, we provide for some brief copula theory. Last, portfolio optimization method, which is useful for our studies in Chapters 3, 4 and 5.

2.1 Capital Asset Pricing Model with Belief functions

2.1.1 Maximum likelihood estimation of capital asset pricing model

The CAPM represents a positive and linear relationship between asset return and systematic risk relative the overall market. The linear regression model is defined as

$$E(R_i) - r_f = \alpha + \beta E(R_M - r_f), \tag{2.1}$$

where $E(R_i)$ is the expected return of the asset, $E(R_M)$ is the expected market portfolio return, r_f is the risk free rate, α is the intercept and β is the equity beta, representing market risk. The observed the historical returns of stock $R_i = (r_{i1}, \dots, r_{in})$ and returns from market $R_M = (r_{m1}, \dots, r_{mn})$. The estimator of beta is a measure of risk for financial analysis and also for risk and portfolio managers. To measure the systematic risk of each stock via beta takes form as

$$\beta = \frac{cov(R_i, R_M)}{\sigma_M^2}\ , \tag{2.2}$$

where σ_M^2 represents the variance of the expected market return. Given that, the CAPM predicts portfolio's expected return should be about the risk and the market returns. The parameter β estimation procedure is defined by Arellano-Valle et al. (2010). The linear regression equation given in (2.1), which has extended into equation as follows:

$$r_i - r_f = \alpha + \beta(r_m - r_f) + \varepsilon_i \tag{2.3}$$

or

$$y_i = \alpha + \beta x_i + \varepsilon_i\,, \tag{2.4}$$

where r_i denotes the return of stock i, r_m is the market return and r_f corresponds to the is free return , so that

$$y_i = r_i - r_f\,, \tag{2.5}$$

and

$$x_i = r_m - r_f, \tag{2.6}$$

represent the return of an asset in excess of risk free rate and the excess return of the market portfolio of assets.

The estimation method with the considering in the financial model is based on the least squares theory under the assumption of the random errors $\varepsilon_1, \dots \varepsilon_n$ are independent and identically distributed according to the normal distribution. The normal density function can be expressed as

$$N(\varepsilon_i, 0, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{\frac{-1}{2\sigma^2}(y - x\beta)^2\}. \tag{2.7}$$

The likelihood function is given by

$$L = \Pi_{n=1}^{n} N(y_i, x_i, \beta, \sigma^2) = (2\pi\sigma^2)^{\frac{-n}{2}} \exp\left\{\frac{-1}{2\sigma^2}(y - x\beta)'(y - x\beta)\right\}. \tag{2.8}$$

2.2 Belief Function

The theory of belief function is a formalism for reasoning with the uncertain, inaccurate and incomplete information. The model comprises several functions including Bel (degree of belief), Dis (degree of disbelief), Unc (degree of uncertainty) and Pls (degree of plausibilty), in range of [0,1]. Belief function can be defined on finite set and infinite set. Let us begin with finite case.

2.2.1 Belief functions on finite set

In the formalism of belief functions, we assign probabilities to sets (see, Pearl (1990)).The belief model as given below (see, Frikha (2014), Liu et al. (2014), Nampak et al. (2014)).

Let Θ be a finite set. Θ is called frame of discernment of the problem of consideration. The power set of Θ is denoted by 2^{Θ}. A basic probability assignment (BPA) is a function $m(.)$ from 2^{Θ} to [0,1] that assigns a number [0,1] to each subset A of Θ. The quantity $m(A)$, called the mass of its subsets. This function satisfies the following conditions:

$$0 \le m(A) \le 1, m(\emptyset) = 0, \sum_{A \subseteq \Theta} m(A) = 1. \tag{2.9}$$

When $m(A) \ge 0, A$ is called focal set of m.To each BPA, we can associate a belief function and a plausibility function are a mapping $Bel(A): 2^{\Theta} \longrightarrow$ [0,1] and $Pl(A) : 2^{\Theta} \longrightarrow$ [0,1] respectively, defined as:

$$Bel(A) = \sum_{B \subseteq A} m(B), \tag{2.10}$$

$$Pl(A) = \sum_{A \cap B \neq \emptyset} m(B), \tag{2.11}$$

where $Bel(A)$ measures the total belief completely attributed to $A \subseteq \Theta$. It is interpreted as the lower bound of probability of A. $Pl(A)$ is interpreted as the upper bound of probability of A.

The two functions satisfied the following properties:

$$Bel(A) \leq Pl(A), \tag{2.12}$$

$$Pl(A) = 1 - Bel(\bar{A}), \tag{2.13}$$

where A is the complement of A and $Bel\bar{A}$ is called a degree of disbelief in A. The uncertainty can be measured by plausibility and belief functions. Eq. (2.14) represents the difference between belief and plausibility.

$$Pl(A) - Bel(A) = Unc. \tag{2.14}$$

If Unc = 0, then $Pl(A) = Bel(A)$.

The schematic description of the relationship between belief, disbelief and uncertain functions is shown as follows:

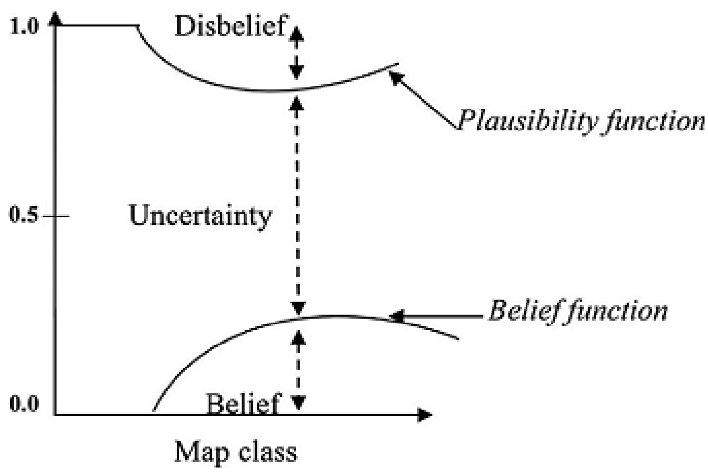

Fig.2.1 Schematic description of the relationship between belief, disbelief and uncertainty (Carranza et al. (2005)).

2.2.2 Belief functions on infinite set

In an infinite case, there may not be a mass function associated with completely monotone function as in the finite case, Denoeux (2013). The definitions are provided which defined by Denoeux (2013) as following:

Let(Ω, B) be a measurable space (i.e., B is a sigma–field, that is a non–empty subset of 2^{Θ}closed under complementation and countable union). A belief function on B is a function $Bel\colon \longrightarrow [0,1]$ verifying the following three conditions:

1. $Bel(\phi) = 0$

2. $Bel(\Omega) = 0$

3. For any $k \geq 2$ and any collection $B_1, \dots, B_k$ of elements of B,

$$Bel(U_{i=1}^{k} B_i) \geq \sum_{\phi \neq I(1,\dots,k)} (-1)^{|I|+1} Bel(\cap_{i \in I} B_i) \tag{2.15}$$

Furthermore, a belief function Bel on (Ω, B) is continuous if for any decreasing sequence $B_1 \supset B_2 \supset B_3 \dots$ of elements of B,

$$\lim_{i \longrightarrow \infty} Bel(B_i) = Bel(\cap_{i \in I} B_i) \tag{2.16}$$

2.2.3 Likelihood–based belief functions

The likelihood-based belief functions have been derived by Shafer (1976). They have been applied by Abdallah et al. (2014), among others, and justified by Denoeux (2014).

Let $x \in X$ be the observable data with a probability density function (pdf) $p_{\theta}X$, where $\theta \in \Theta$ is an unknown parameter. In this study, we use the method proposed by Shafer (1976).The belief function be derived from the Likelihood Principle and Least Commitment Principle (LCP). The information about Q can be represented by the likelihood function which is defined by $L_x(\theta) = p_\theta X$ for all $\theta \in \Theta$. The likelihood ratio is meant to be a "relative plausibility", which can be written as:

$$\frac{pl_x(\theta_1)}{pl_x(\theta_2)} = \frac{L_x(\theta_1)}{L_x(\theta_2)}. \tag{2.17}$$

For all $(\theta_1, \theta_2) \in \Theta^2$ or, equivalent $pl_x(\theta) = cL_x(\theta)$

for all $\theta \subseteq \Theta$ and some positive constant c. From LCP, it can be implied that the highest possible value of C is $\frac{1}{sup_{\theta\in\Theta}} = L\left(\frac{\theta}{X}\right)$. Thus, the contour function is defined as follow:

$$pl(\theta; x) = \frac{L(\theta; x)}{sup_{\theta\in\Theta} L(\theta; x)}. \tag{2.18}$$

The information about θ are expressed by the belief function Bel_A^{Θ} with contour function pl_x , i.e., with corresponding plausibility function $Bel_A^{\Theta} = sup_{\theta\in A} pl_x(A)$, for all $A \subseteq \Theta$.The focal sets of Bel_A^{Θ} the levels sets of pl_x defined as follows:

$$\Gamma_x(\omega) = \{\theta \in \Theta | pl_x(\theta) \geq \omega\} \tag{2.19}$$

for $\theta \in [0,1]$. Equation (2.19) is called plausibility regions. With the inducing of the Lebesgue measure λ on [0,1] and multi–valued mapping Γ_x from $[0,1] \longrightarrow \Theta^2$ the belief function is equivalent to the random set (see, Kanjanatarakul et al. (2014)). We remark that the MLE of θ is the value of θ with highest plausibility.

2.2.4 Incorporating the belief functions

The objective is to forecast the risk premium of the return of stock i, $y_i = r_i - r_f$. The methodology to incorporate the belief function framework into the prediction procedure follows Kanjanatarakul et al. (2014). From the CAPM equation above, the return equation can be written as:

$$y_i = \alpha + \beta x + \sigma F^{-1}(u), \tag{2.20}$$

where $F \sim Normal(0,1)$ and $U \sim Uniform(0,1)$.

Discussed in Kanjanatarakul et al. (2014), the forecasting problem is the inverse problem of the regular inference problem. Given the knowledge on the set of parameters $\theta = (\alpha, \beta, \sigma)$ and the distribution $F(.)$, the future value of y_i can be forecasted.

Belief function framework allows us to forecast an interval $[y_i^L, y_i^U]$ for the future value of y_i. The estimation of $[y_i^L, y_i^U]$ can be done using Monte Carlo method. Given a set two Uniform$(0,1)$ random variables(u_s, ω_s), in each simulation s, the lower bound $y_{i,s}^L$ and the upper bound $y_{i,s}^U$ solve the following optimization problems respectively,

$$y_{i,s}^L = \min_{\theta} \alpha + \beta x + \sigma F^{-1}(u_s), \tag{2.21}$$

subject to

$$pl(\theta) \geq \omega_s, \tag{2.22}$$

and

$$y_{i,s}^U = \max_{\theta} \alpha + \beta x + \sigma F^{-1}(u_s), \tag{2.23}$$

subject to

$$pl(\theta) \geq \omega_s. \tag{2.24}$$

In the constraints, the plausibility function $pl(\theta, u_s)$ can be derived from the likelihood function. Therefore, using the likelihood function, the plausibility function is as follows:

$$pl(\theta) = \frac{L(\theta)}{L(\theta^*)}, \tag{2.25}$$

where θ^* is such that $L(\theta^*) \geq L(\theta), \forall \theta$.

The belief and the plausibility functions corresponding to a given set A can be calculated by:

$$\widehat{bel}(A) = \frac{1}{N}\#\{s \in \{1, \dots, N\} | [y_{i,s}^{L}, y_{i,s}^{U}] \subset A\}, \tag{2.26}$$

$$\widehat{pl}(A) = \frac{1}{N}\#\{s \in \{1, \dots, N\} | [y_{i,s}^{L}, y_{i,s}^{U}] \cap A \neq \emptyset\}. \tag{2.27}$$

The lower bound and the upper bound of the prediction for y_i is, thus,

$$\hat{y}_i^{L} = E(y_{i,s}^{L}) = \frac{1}{N}\sum y_{i,s}^{L}, \tag{2.28}$$

$$\hat{y}_i^{U} = E(y_{i,s}^{U}) = \frac{1}{N}\sum y_{i,s}^{U}. \tag{2.29}$$

2.3 Interval –Valued Data

2.3.1 An interval-valued data in a linear regression model

Suppose we can observe an i.i.d. random paired intervals variables $x_i = \underline{x_i}, \overline{x_i}$ and $y_i = [\underline{y_i}, \overline{y_i}], i = 1,2, \dots, n$ where $\overline{x_i}, \overline{y_i}$ are the maximum values of x_i and y_i and $\underline{x_i}, \underline{y_i}$, are the minimum values of x_i and y_i. Additionally, we can rewrite the value of x_i, y_i in the form of intervals as

$$x_i = [x_i^m - x_i^r, x_i^m + x_i^r], \tag{2.30a}$$

$$y_i = [y_i^m - y_i^r, y_i^m + y_i^r], i = 1,2, \dots, n, \tag{2.31b}$$

where x_i^m, y_i^m is the mid-points of x_i and $y_i$and x_i^r, y_i^r is the radii of x_i and y_i, satisfying $x_i^r, y_i^r \geq 0$ Suppose, we consider the following linear regression model given by

$$y_i = ax_i + b + \varepsilon_i, i = 1,2, \dots, n. \tag{2.32}$$

Analogously, it is easily to interpret the meaning of x_i, y_i by the distance of centers and radii as the following equations:

$$x_i = \tilde{x}_i + \Delta x_i, x_i \in N(0, k_0 \Delta x_i), \tag{2.33a}$$

$$y_i = \tilde{y}_i + \Delta y_i, y_i \in N(0, k_0 \Delta y_i), \tag{2.33b}$$

where $\tilde{x}_i, \tilde{y}_i$ are the centers of $x_i$and y_i, respectively. Then, $\Delta x_i = \frac{\overline{x_i} - x_i}{2}$, $\Delta y_i = \frac{\overline{y_i} - y_i}{2}$ are the radii of x_i and y_i, respectively and $m_{xi} = \frac{\overline{x_i} + x_i}{2}$, $m_{yi} = \frac{\overline{y_i} + y_i}{2}$ are the mid-point of x_i and y_i, respectively. Thus, given the linear regression for the interval-valued data we have

$$\tilde{y}_i + \delta y_i = a\tilde{x}_i + a\delta x_i + b, \tag{2.34a}$$

$$\tilde{y}_i = a\tilde{x}_i + b + (a\delta x_i - \delta y_i), \tag{2.34b}$$

where $(a\delta x_i - \delta y_i) \sim N(0,1) \equiv N\left(0, k_0\sqrt{a^2}\Delta x_i^2 + \Delta y_i^2\right)$. Assume that $a\delta x_i - \delta y_i$ is an i.i.d thus, we can estimate parameter a, b, k_0 by the maximum likelihood function given by

$$\max_{a,b,k_0} \ln L(a, b, k_0 \mid \left(\left[\underline{x_i}, \overline{x_i}\right], \left[\underline{y_i}, \overline{y_i}\right]\right), i = 1, \dots, n)$$
$$= \max_{a,b,k_0}\{\sum_{i=1}^{n} ln\emptyset(\frac{\tilde{y}_i - a\tilde{x}_i - b}{k_0\sqrt{a^2}\Delta x_i^2 + \Delta y_i^2})\}, \tag{2.35}$$

where $\emptyset(.) \sim N(0,1)$. This approach was already developed in Sun and Li (2015).

2.3.2 Goodness of fit in linear regression model for an interval-valued data

In the deterministic linear regression model, we use variance to describe variation of the variable interested and so that as we knew the ratio $\frac{a^2Var(X)}{Var(Y)} \in [0,1]$ can be explained as an indication of goodness-of-fit. In this study, we used the concept of the chi-squared test (χ^2) of the goodness of fit. Recall that $\sigma_{xi} = k_0\Delta x_i$ and $\sigma_{yi} = k_0\Delta y_i$. Given the simple linear regression, we have

$$y_i = ax_i + b, \tag{2.36a}$$

$$y_i^m + \delta_{yi} = ax_i^m + a\delta_{xi}, \tag{2.36b}$$

$$y_i^m - ax_i^m - b = a\delta_{xi} - \delta_{yi}, \tag{2.36c}$$

where $\delta_{xi}, \delta_{yi} \sim N(0, \sigma^2)$. Thus, we have $a^2\delta_{xi}^2 + \delta_{yi}^2$, by replacing $k_0^2(a^2\Delta x_i^2 + \Delta y_i^2)$ to above equation (2.24). The empirical chi-squared test (χ^2) is obtained by estimated this following equation

$$\chi_{cal}^2 = \sum_{i=1}^{n} \frac{(y_i^m - ax_i^m - b)^2}{k_0^2(a^2\Delta x_i^2 + \Delta y_i^2)}, \tag{2.37}$$

where the degree of freedom is $n-2$.

2.3.3 Beta estimation with interval data

From the CAPM model in equation (2.1), we calculate the β coefficient through the likelihood by equation (2.35) instead. Suppose we have observed the realization of interval stock return (R_A), $[\overline{R_{Ai}}, \underline{R_{Ai}},] = \left[\left(\overline{r_{a1}}, \underline{r_{a1}}, \dots, \overline{r_{an}}, \underline{r_{an}}\right)\right], i = 1,2,\dots,n$ and interval return from market $[\overline{R_{Mi}}, \underline{R_{Mi}},] = \left[\left(\overline{r_{m1}}, \underline{r_{m1}}, \dots, \overline{r_{mn}}, \underline{r_{mn}}\right)\right], i = 1,2,\dots,n$ over the past N years. These observations will be assumed an independent random. From likelihood for an interval values we have

$$\max_{a,b,k_0} L\left(a, b, k_0 \mid \left(\left[\underline{R_{Mi}}, \overline{R_{Mi}}\right], \left[\underline{R_{Ai}}, \overline{R_{Ai}}\right]\right), i = 1, \dots, n\right).$$

$$= \max_{a,b,k_0} \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi k_0^2(a^2\Delta Rm_i^2 + \Delta Ra_i^2)}} \exp\left[-\frac{1}{2}\frac{(Ra_i^m + aRm_i^m - b)^2}{k_0^2(a^2\Delta Rm_i^2 + \Delta Ra_i^2)}\right]\right. \tag{2.38}$$

2.4 Copula

Copula is a powerful tool to model the non-normal distribution. It can capture the complicated correlation between variables, including linear or non-linear. In finance and econometrics, Copula approach has been popular for modeling nonlinear stochastic relationships between two or more variables. Generally, Pearson correlation is used for variables dependence measure in linear relation and based on the assumption of normality. However, the financial returns tend to show asymmetric dependence. Pearson correlation can not capture the non-linear relations between variables and it is not invariant under strictly increasing transformations (Schirmacher and Schirmacher, 2008).

Mathematically, copula is a multivariate distribution whose one-dimensional margins are uniform on the interval [0, 1].The definition of copula as follows:

Definition: $C\colon [0,1]^n \to [0,1]$ is a n-dimensional copula if C is a joint cumulative distribution function of a n-dimensional random vector on the unit cube $[0,1]^n$ with uniform marginal.

In analytic terms, $C\colon [0,1]^n \to [0,1]$ is a n-dimensional copula if

1.$C(u_1, \dots, u_{j-1}, 0, u_{j+1}, \dots, u_n) = 0$, the copula is zero if one of the arguments is zero,

2. $(1, \dots, 1, u, 1, \dots, 1) = u$, the copula is equal u if one of the arguments is u and all others 1,

Properties 1 and 2 ensure that marginal distributions are uniform distributions.

3. C is n-non-decreasing, i.e., for each hyper rectangle $B = \prod_{i=1}^{n} [x_i, y_i] \subseteq [0,1]^n$ the C volume of B is non-negative:

$$\int_B dC(u) = \sum_{z \in \times_{i=1}^{n} \{x_i, y_i\}} (-1)^{N(z)} C(z) \geq 0$$

where the $N(z) = \#\{k | z_k = x_k\}$.

For ensures that copula is a proper cumulative distribution function.

Suppose we have a set of observations (x_1, y_1) , $(x_2, y_2), \dots, (x_n, y_n)$ from an unknown bivariate distribution $H(x, y)$. According to Sklar's theorem (1959), the joint distribution function $H(x, y)$ of any pair of continuous random variables (x, y) be written in the form

$$H(x, y) = C(F(x), G(y)) \ \ x, y \in \mathbb{R}, \tag{2.39}$$

where $F(x)$and *G(y)* are the marginal distributions of *X* and *Y*, and *C* is a function mapping [0, 1] × [0, 1] → [0, 1] known as a copula.

The theorem is valid in the bivariate case (n = 2) and in all higher dimensions (n > 2). Consider in multivariate distribution. Let *F* be an *n*-dimensional distribution function with marginal functions $F_1, F_2, \dots, F_n$. Then there exists an *n*-dimensional copula *C* such that for all $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$,

$$F(x_1, x_2, \dots, x_n) = C(F_1(x_1), F_2(x_2), \dots, F_n(x_n)). \tag{2.40}$$

If the marginal functions $F_1, F_2, \dots, F_n$ are all continuous, then *C* exists and is unique.

The inverse function of $F_i (i = 1,2, \dots, n)$ respectively are F_i^{-1} $(i = 1,2, \dots, n)$, setting $u_i = F_i (i = 1,2, \dots, n)$, whose copula function can be calculated as :

$$C(u_1, u_2, \dots, u_n) = F(F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_n^{-1}(u_n)), \tag{2.41}$$

where $u_i = F_i (i = 1,2, \dots, n)$ are the probability integral transformation (PIT) of the marginal models.

The reverse of (2.40) where any multivariate distribution *F* can be written in terms of its marginals using a copula representation.

Assume F_i and C are differentiable. The joint density function $f(x_1, x_2, \dots, x_n)$ is defined as:

$$f(x_1, x_2, \ldots, x_n) \\ = c(F_1(x_1), F_2(x_2), \ldots, F_n(x_n))f_1(x_1), f_2(x_2) \ldots f_n(x_n), \quad 2.42)$$

where the density of F_i is given by $f_i(x_i)$ and the density of the copula is given by:

$$c(u_1, u_2, \ldots, u_n) = \frac{\partial^n C(u_1, u_2, \ldots, u_n)}{\partial u_1 \partial u_2, \ldots, \partial u_n}. \quad (2.43)$$

2.4.1 Pair-copula decomposition of multivariate distribution

For high-dimension distribution, multivariate copulas are obtained by pair-copula construction method (PCC), was proposed by Aas et al. (2009). In PCC models, bivariate copulas are used as building blocks. Consider a vector $X = (X_1, \ldots, X_n)$ of random variables with density function $f(x_1, \ldots, x_n)$. This density can be factorized as

$$f(x_1, \ldots, x_n) \\ = f_n(x_n). f(x_{n-1}|x_n). f(x_{n-2}|x_{n-1}, x_n) \ldots f(x_1|x_2, \ldots, x_n). \quad (2.44)$$

De Melo Mendes (2010) illustrated that the conditional densities in (2.44) can be written as functions of corresponding copula densities by using the general formula as

$$f(x\,|v_1, \ldots, v_n) = c_{xv_j|v_{-j}}\{F(x|v_{-j}), F(v_j|v_{-j})\}. f(x|v_j), \quad (2.45)$$

where v_{-j} denotes the n dimensional vector v excluding the j th component.$c_{xv_j|v_{-j}}(.,.)$ is a bivariate marginal copula density.

The conditional densities in (2.44) by means of (2.45) we derive a decomposition for $f(x_1, \ldots, x_n)$ that only consists of univariate marginal distributions and bivariate copulas. Thus we obtain the pair-copula decomposition for the *n*-dimensional copula $c_1, \ldots, c_n$, a factorization of a *n*-dimensional copula based only in bivariate copulas.

The pair-copula construction involves marginal conditional distributions of the form $F(x\,|v)$. For every j, Joe (1996) showed that

$$F(x\,|v) = \frac{\partial C_{xv_j|v_{-j}}\{F(x|v_{-j}), F(v_j|v_{-j})\}}{\partial F(v_j|v_{-j})}. \tag{2.46}$$

2.4.2 Vine Copula

Vine copula are a class of multivariate dependence model that are constructed on the theory of PCC models. Vine copula are more flexible than the bivariate copulas because they allow for decompositions. In addition, Vine copula can capture the asymmetry model as well as the tail behavior of the underlying risk exposures in the context of multivariate distribution (Han et al., 2014).

Vines copula are dependence models that the multivariate distribution function be decomposed into bivariate copulas and marginal densities. The term vine was used because the shape of induced dependence structure can be seen like a grape vine. Three components of vine copula are the tree structure, the copula family for each edge in the tree structure and the corresponding dependence parameters for each pair copula (Czado et al., 2013).

The graphical of these structure are nested trees and then it is called regular vines. A vine on n variables is a nested set of trees $T_1, \ldots T_{n-1}$, where the edges of tree j are the nodes of the tree $j+1$ with $j = 1, \ldots, n-2$. Two subclasses of regular vines are canonical (C-Vine) and drawable vine (D-vine) in which two edges in tree j are joined by an edge in tree $j+1$ only if these edges share a common node. The n-dimensional density function of C-vine and D- vine are defined as follows:

C- vine:

$$f(x) = \prod_{k=1}^{n} f_k(x_k) . \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{j,\mathrm{j+i}|1,2,\dots j-1}(F(x_j / x_1, \dots, x_{j-1}), F(x_{j+i} / x_1, \dots, x_{j-1}). \quad (2.47)$$

D-vine:

$$f(x) = \prod_{k=1}^{n} f_k(x_k) . \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{i,i+\mathrm{j}|\mathrm{i+1},2,\dots i+j-1}(F(x_i / x_{i+1}, \dots, x_{i+j-1}), F(x_{i+j} / x_{i+1}, \dots, x_{i+j-1}). \quad (2.48)$$

where f_i is the marginal density of x_i , $c_{i,j|k}$ is the bivariate copula distribution function. *j* refers to tree, i refers to the edge of a tree. Only one node is connected with *m-j* edges and other nodes connected with one edge in trees of C-vine respectively, and trees of D-vine just as lines.

2.4.3 Inference for a C-Vine and D-Vine

The parameters of the C-Vine density given by (2.47) or D-Vine density given (2.48) can be estimated by maximum (log) likelihood estimation (MLE).

For the C-Vine, Log-likelihood is given by

$$l_{CV}(\theta|x) = \sum_{j=1}^{n-1} \sum_{i=1}^{n-j} \sum_{t=1}^{T} \log[c_{j,j+i|1,\dots,j-1}\{F(x_{j,t}|x_{1,t}, \dots, x_{j-1,t}), F(x_{j+1,t}|x_{1,t}, \dots x_{j-1,t})\}. \quad (2.49)$$

For the D-Vine, Log-likelihood is given by

$$l_{DV}(\theta|x) = \sum_{j=1}^{n-1}\sum_{i=1}^{n-j}\sum_{t=1}^{T}\log[c_{i,i+j|i+1,..,i+j-1}\{F(x_{i,t}|x_{i+1,t}\,,\ldots,x_{i+j-1,t}),F(x_{i+1,t}|x_{i+1,t}\,,\ldots x_{i+j-1,t})\}], \tag{2.50}$$

where θ is the parameter vector , $c_{j,j+i|1,\ldots,j-1}$ and $c_{i,i+j|i+1,..,i+j-1}$ are the bivariate copula density in equation (2.44) and (2.45).

2.4.4 Canonical Vine (C-vine)

A C-vine is a regular vine such that each tree $T_j$has a unique node of degreed-j.The node with maximal degree in T_1 is the root, Here, we focus on eight assets, for eight-dimension (n=8) C-vine copulas can written as

T_1: Nodes: 1(root), 2,3,4,5,6,7,8; Edges: 12,13,14,15,16,17,18
T_2: Nodes: 12(root)13,14,15,16,17,18; Edges: 23|1,24|1,25|1,26|1,27|1,28|1
T_3: Nodes: 23|1(root)24|1,25|1,26|1,27|1,28|1; Edges: 34|12,35|12,36,12,37|12,38|12
T_4: Nodes: 34|12,35|12,36|12,37|12,38|12; Edges; 45|123,46|123,47|123,48|123
T_5: Nodes: 45|123,46|123,47|123,48|123; Edges; 56|1234,57|1234,58|1234
T_6: Nodes: 56|1234,571234,58|1234; Edges; 67|12345,68|12345
T_7: Nodes: 67|12345,68|12345; Edges; 78|123456

The decomposition of joint densities in terms of C-vines copulas is illustrated as follows.

Consider n = 8. As in the case of C-copulas, consider a decomposition of the following form

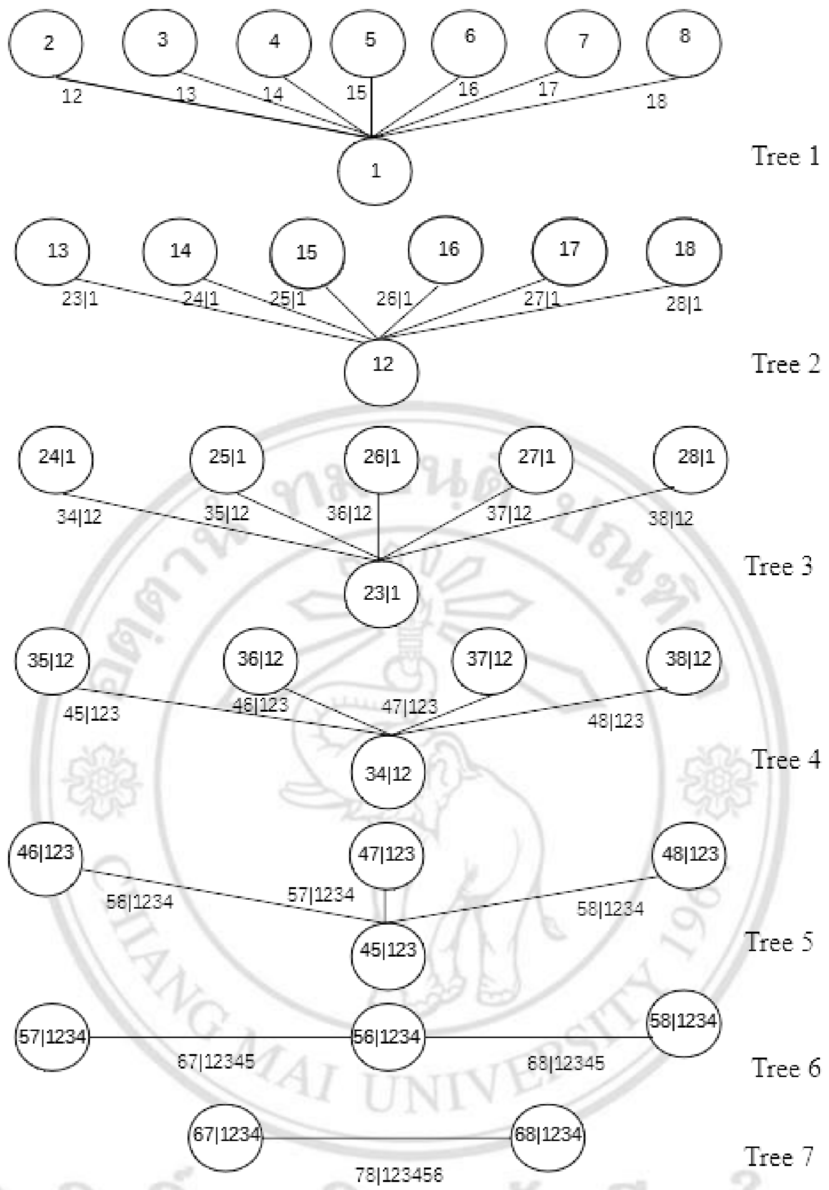

Fig 2.2 The C-vine structure for eight dimensional variables.

The decomposition of joint densities in terms of C-vines copulas is illustrated as follows.

$$f(x_1, x_2, \dots, x_8) = \prod_{i=1}^{8} f_i(x_i).\, c_{12}(F_1.F_2).\, c_{13}(F_1.F_3).\, c_{14}(F_1.F_4).\, c_{15}(F_1.F_5)$$

$$.\, c_{16}(F_1.F_6).\, c_{17}(F_1.F_7).\, c_{18}(F_1.F_8) c_{23|1}(F_{2|1}.F_{3|1})$$

$$.\, c_{24|1}(F_{2|1}.F_{4|1}).\, c_{25|1}(F_{2|1}.F_{5|1}).\, c_{26|1}(F_{2|1}.F_{6|1})$$

$$.\, c_{27|1}(F_{2|1}.F_{7|1}).\, c_{28|1}(F_{2|1}.F_{8|1}).\, c_{34|12}(F_{3|12}.F_{4|12})$$

$$.\, c_{35|12}(F_{3|12}.F_{5|12}).\, c_{36|12}(F_{3|12}.F_{6|12}).\, c_{37|12}(F_{3|12}.F_{7|12})$$

$$.\, c_{38|12}(F_{3|12}.F_{8|12}).\, c_{45|123}(F_{4|123}.F_{5|123}) c_{46|123}(F_{4|123}.F_{6|123})$$

$$.\, c_{47|123}(F_{4|123}.F_{7|123}).\, c_{48|123}(F_{4|123}.F_{8|123})$$

$$.\, c_{56|1234}(F_{5|1234}.F_{6|1234}).\, c_{57|1234}(F_{5|1234}.F_{7|1234})$$

$$.\, c_{58|1234}(F_{5|1234}.F_{8|1234}).\, c_{67|12345}(F_{6|12345}.F_{7|12345})$$

$$.\, c_{68|12345}(F_{6|12345}.F_{8|12345}).\, c_{78|123456}(F_{7|123456}.F_{8|123456}).$$

2.4.5 Drawable Vine (D-vine)

The decomposition of the joint density in terms of bivariate pairwise copulas and marginals is drawable, and hence is called a D-vine. With this drawable vine copula, the joint density is obtained simply by multiplying all (bivariate) copula densities which appeared in the tree together with all marginal densities. The usefulness of graphical displays is this. When trying to model dependencies in a multivariate model (i.e., we do not know the joint distribution), we choose a D-vine, according to important pairwise dependencies of interest, and from which we have a "formula" to arrive at the joint distribution, i.e., to arrive at a model capturing the dependencies of interest. How to use D-vine copulas to build multivariate models? In general, we should figure out that, any n-dimensional copula density can be decomposed in $\frac{\boldsymbol{n(n-1)}}{\boldsymbol{2}}$different ways.

$n = 8, X = X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8$ a possible D-vine is

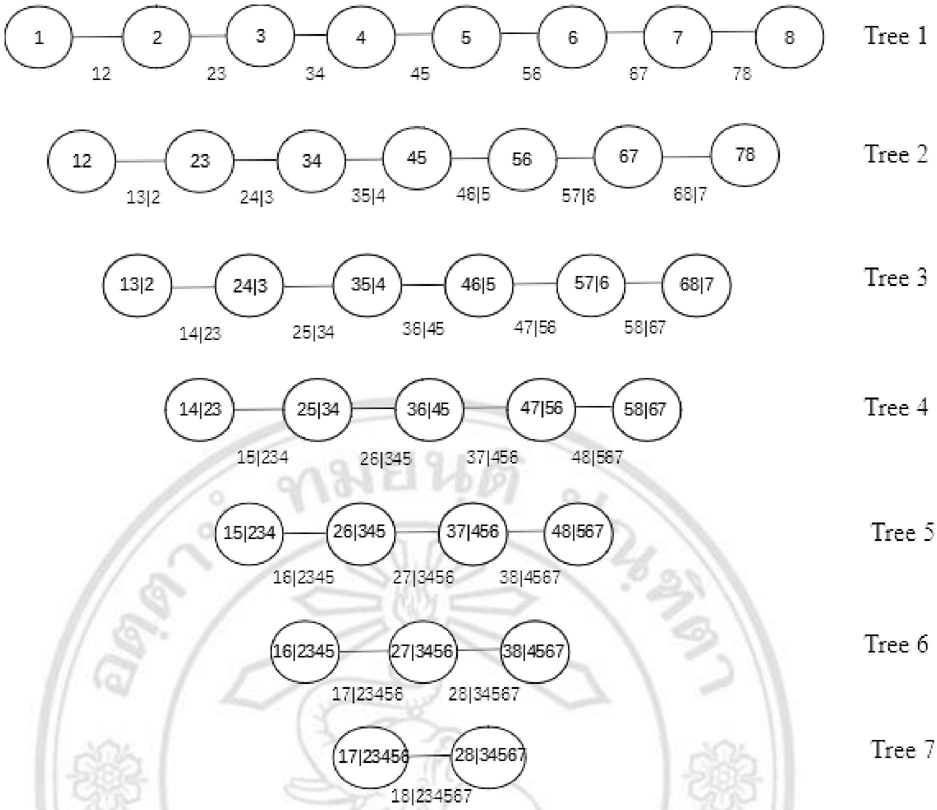

Fig.2.3 The D-vine structure for eight dimensional variables.

The multivariate(density) model is as follow:

$$f(x_1, x_2, \dots, x_8) = \prod_{i=1}^{8} f_i(x_i).\, c_{12}c_{23}\, c_{34}c_{45}c_{56}c_{67}c_{78}$$

$$\cdot\, c_{13|2}c_{24|3}c_{35|4}c_{46|5}c_{57|6}c_{68|7}$$

$$\cdot\, c_{14|23}c_{25|34}c_{36|45}c_{47|56}c_{58|67}$$

$$\cdot\, c_{15|234}c_{26|345}c_{37|456}c_{48|567}$$

$$\cdot c_{16|2346}c_{27|3456}c_{38|4567}$$

$$\cdot c_{17|23456}c_{28|34567}$$

$$\cdot\ c_{18|234567}$$

2.5 Value at Risk (VaR) and Conditional Value at Risk (CVaR)

Value at Risk (VaR) is defined as the possible maximum loss over a given holding period within a fixed confidence level (cl). Following Sarykalin et al. (2008), mathematical the definition of VaR:

Definition VaR : Let X is a random variable with the cumulative distribution function $F_X(z) = P(X \leq z)$. X has meaning the of loss. Given confidence level $\alpha \in (0,1)$ the VaR of X is

$$VaR_\alpha(X) = min\{z|F_X(z) \geq \alpha\}. \tag{2.51}$$

Artzner et al. (1998) proposed a set of axioms for evaluation of risk measure. Any risk measure satisfies these axioms is said to be coherent. The four axioms are *Monotonicity*, *Translation equivalence*, *Sub-additivity*, *and Positive Homogeneity*. However, VaR based on the assumption of a systematical return distribution and lacks of subadditivity axiom. Therefore, VaR is not qualify for coherent risk measure.

Conditional value at Risk (CVaR) was introduced by Rockafellar and Uryasev (2000). CVaR is more appealing than VaR because it can measure risk of fat tails and asymmetric distributions of losses and satisfies sub-additivity axiom.

VaR and CVaR are related in connections. For $\alpha\epsilon(0,1)$, the inverse of cumulative distribution function of $F_X(z)$ is the α −quantile of X is the smallest value $F_X^{-1}(\alpha)$ as follow:

$$F_X^{-1}(\alpha) = \inf\{x|F(x) \geq \alpha\}, \tag{2.52}$$

and denoted a risk measure, known as

$$VaR_\alpha(X) = F_X^{-1}(\alpha). \tag{2.53}$$

CvaR is the conditional expectation of losses that exceed the $VaR_\alpha(X)$level. For random variables with continuous distribution functions, $CVaR_\alpha(X)$ equals the conditional expectation of X subject to$X \geq VaR_\alpha(X)$.Given confidence level $\in [0,1]$, the CVaR of X is the mean of the generalized α-tail distribution :

$$CVaR_\alpha(X) = \int_{-\infty}^{\infty} z dF_X^\alpha(z), \tag{2.54}$$

where

$$F_X^\alpha(z) = \begin{cases} 0 & when \quad z < VaR_\alpha(X) \\ \dfrac{F_X(z) - \alpha}{1-\alpha} & when \quad z \geq VaR_\alpha(X). \end{cases} \tag{2.55}$$

The relationship between VaR and CVaR is illustrated in the following graph:

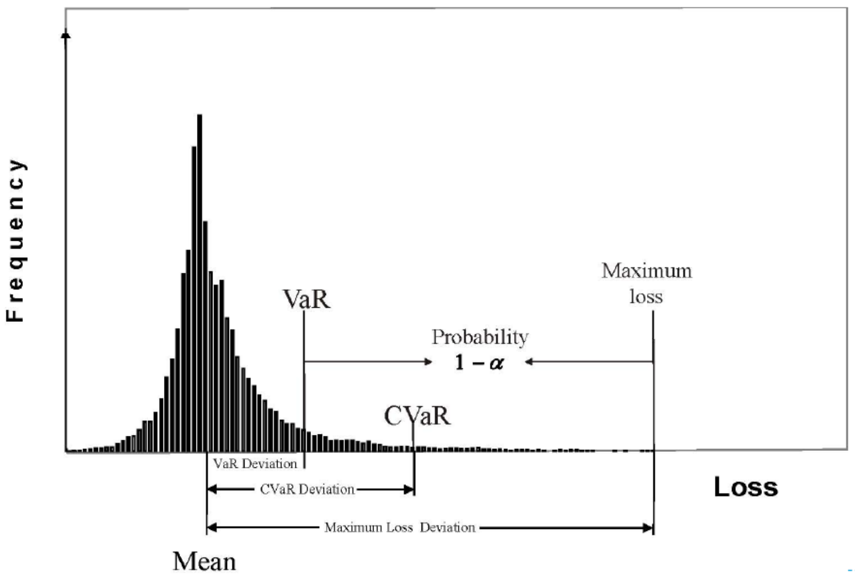

Fig 2.4 Graphical representation of VaR and CVaR (Sarykalin et al. (2008)).

2.6 CVaR Optimization

Portfolio optimization problem is the process of choosing the proportions of variety assets to be held in portfolio under the constraints. The purpose is to make investment maximizing returns to investors. The CVaR measure is used in the portfolio optimization problem not only because it is a coherent measure of risk but also, it is more in tune with the loss function of the tail distribution.

2.6.1 Optimal Portfolio with Conditional Value at Risk via Vine-Copulas

We start our calculation of VaR and CVaR of an equally weighted portfolio and then, the optimal portfolio can be constructed by minimizing CVAR subject to maximum returns. The procedure of optimization, we refer to the paper from Autchariyapanitkul (2014). The following formula can show as below:

$$Min\ CVaR = E[r_p | r \leq r_\alpha] \quad (2.56a)$$

Subject to

$$E(r_p) = w_1(r_1) + w_2(r_2) + \cdots + w_n(r_n) \quad (2.56b)$$

$$w_1 + w_2 + \cdots + w_n = 1 \quad (2.56c)$$

$$0 \leq w_i \leq 1, where\ i = 1,2, \ldots, n$$

where r_α is the lower $\alpha - quantile$, and r_p is the return on individual asset at time *t*. w_i is the weight of returns.

We use vine copulas to extract dependence structure between CAPM equations and then use the solutions of C-vine and D-vine copulas parameters to create an efficient portfolio and find the optimal solutions for the expected returns with minimum lost.

Next, we stimulate the errors terms of each stocks to get the simulated return through CAPM. By generating a uniform marginals from Vine Copulas and transform those marginal to be an errors using quantiles function of normal distribution. In this study, we stimulate 1,000,000 samples for each stock get the return of portfolio with $r = \sum_{i=1}^{n} W_i X_i$, where W_i is the weight of returns. X_i is the individual return. Finally, we compute the portfolio's VaR and CVaR.