CHAPTER 1

Introduction

Measures of complete dependence are tools to quantify dependence among random vectors.

In 2010, Siburg and Stoimenov [1] defined a measure of dependence

$$\omega(X,Y) = \sqrt{3\int_{[0,1]^2}\left[\left(\frac{\partial}{\partial x}C_{X,Y}(x,y)-y\right)^2+\left(\frac{\partial}{\partial y}C_{X,Y}(x,y)-x\right)^2\right]\mathrm{d}x\mathrm{d}y}$$

$$= \sqrt{3\int_{[0,1]^2}\left[\left(\frac{\partial}{\partial x}C_{X,Y}(x,y)\right)^2+\left(\frac{\partial}{\partial y}C_{X,Y}(x,y)\right)^2\right]\mathrm{d}x\mathrm{d}y-2}$$

for any continuous random variables X and Y. Dette et al. [2] later defined a new measure

$$\begin{aligned} r(X,Y) &= 6\int\left|\frac{\partial}{\partial x}C_{X,Y}(x,y)-y\right|^2\mathrm{d}x\mathrm{d}y \\ &= 6\int\left(\frac{\partial}{\partial x}C_{X,Y}(x,y)\right)^2\mathrm{d}x\mathrm{d}y-2 \end{aligned}$$

which can be considered as an asymmetric version of $\omega(X,Y)$, that is,

$$\omega(X,Y) = \sqrt{\frac{r(X,Y)+r(Y,X)}{2}}.$$

During the same period, Trutschnig [3] also defined a measure of complete dependence ζ_1 based on the Markov operator associated with X and Y. When written using the copula $C_{X,Y}$ associated with X and Y,

$$\zeta_1(Y|X) = 3\int\int\left|\frac{\partial}{\partial x}C_{X,Y}(x,y)-y\right|\mathrm{d}x\mathrm{d}y.$$

Both r and ζ_1 are similar in the sense that both measures reach the maximum value if and only if the random variable Y is a measurable function of the random variable X.

Recently, Tasena and Dhompongsa [4] also defined measures of complete dependence

$$\omega_k(Y|X) = \left[\int\int\left|F_{Y|X}(\vec{y}|\vec{x})-\frac{1}{2}\right|^k\mathrm{d}F_X(\vec{x})\mathrm{d}F_Y(\vec{y})\right]^{\frac{1}{k}}$$

and

$$\overline{\omega}_k(Y|X) = \left[\frac{\omega_k^k(Y|X)-\hat{\omega}_k(Y)}{\omega_k^k(Y|Y)-\hat{\omega}_k(Y)}\right]^{\frac{1}{k}}$$

where $\hat{\omega}_k(Y) = \int \left| F_Y(\vec{y}) - \frac{1}{2} \right|^k \mathrm{d}F_Y(\vec{y})$ for any random vectors X and Y. They are able to show that both $\omega_k(Y|X)$ and $\overline{\omega}_k(Y|X)$ reach the maximum value if and only if Y is a measurable function of X. Moreover,

$$\overline{\omega}_2(Y|X) = \sqrt{\frac{\int\int \left(F_{Y|X}(\vec{y}|\vec{x}) - F_Y(\vec{y})\right)^2 \mathrm{d}F_X(\vec{x})\mathrm{d}F_Y(\vec{y})}{\int F_Y(\vec{y})(1 - F_Y(\vec{y}))\mathrm{d}F_Y(\vec{y})}}$$

which can be considered as a generalization of $r(X, Y)$.

It can be easily seen that both ω_k and $\overline{\omega}_k$ depend on F_Y except only for the case Y is a continuous random variable. Also,

$$\overline{\omega}_k(Y|X) \neq \left[\frac{\int\int \left|F_{Y|X}(\vec{y}|\vec{x}) - F_Y(\vec{y})\right|^k \mathrm{d}F_X(\vec{x})\mathrm{d}F_Y(\vec{y})}{\omega_k^k(Y|Y) - \hat{\omega}_k(Y)}\right]^{\frac{1}{k}}$$

in general.

Let $X = (X_1, ..., X_n)$ be an absolutely continuous random vector and F_X be the distribution function of X. We define a transformation function Ψ_{F_X} from $\mathbb{R}^n$ into $[0,1]^n$ by letting

$$\Psi_{F_X}(x_1, ..., x_n) = \left(F_{X_1}(x_1), F_{X_2|X_1}(x_2|x_1), ..., F_{X_n|(X_1,...,X_{n-1})}(x_n|(x_1, ..., x_{n-1}))\right)$$

for all $(x_1, ..., x_n) \in \mathbb{R}^n$.

It is known that $U = \Psi_{F_X}(X)$ has uniform distribution ([5]).

For any two absolutely continuous random vectors X and Y of dimensions m and n, respectively, we know that random vectors $U = \Psi_{F_X}(X)$ and $V = \Psi_{F_Y}(Y)$ are uniform random vectors. Since the *linkage* is a joint distribution function of uniform random vectors, we obtain that the joint distribution function of U and V is a linkage, and this linkage is called the *linkage associated with random vectors X and Y*. In this thesis, $C_{X,Y}$ denotes the linkage associated with random vectors X and Y.

Let X and Y be two absolutely continuous random vectors of dimensions m and n, respectively. In this thesis, we define the measure of complete dependence of Y given X by

$$\zeta_p(Y|X) = \left[\int_{[0,1]^m}\int_{[0,1]^n} \left|\frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u},\vec{v}) - \Pi(\vec{v})\right|^p \mathrm{d}\vec{u}\mathrm{d}\vec{v}\right]^{\frac{1}{p}}$$

where $C_{X,Y}$ is the linkage associated with X and Y.

This thesis comprises of five chapters. Chapter 1 is an introduction of this thesis. Chapter 2, we list some useful results, definitions and notations that will be used throughout this thesis. Chapter 3, we present the measure of complete dependence for random vectors. Chapter 4, we give calculation examples of this measure. Finally, the conclusion and discussion of this thesis are given in Chapter 5.