CHAPTER 3

Main Results

3.1 Measure of Complete Dependence Based on Conditional Distribution Functions

In this chapter, we define the measure of dependence for random vectors using the conditional distribution functions.

Definition 3.1.1. Let X and Y be random vectors.
The *measure of dependence* φ *of* Y *given* X is defined by

$$\varphi(Y|X) = \int\int \left|F_{Y|X}(v|u) - F_Y(v)\right| \mathrm{d}F_X(u)\mathrm{d}F_Y(v)$$

where $F_{Y|X}$ is the conditional distribution function of Y given X.

We will figure out the maximum value of φ using Lemma 3.1.3. The proove of Lemma 3.1.3 is very complicated. Therefore, we separated parts of its proof into the following lemmas.

Lemma 3.1.1. *Let A be a metric space and μ be a Borel probability measure on A such that $\mu(\overline{\mathcal{B}(x,\epsilon)} \setminus \mathcal{B}(x,\epsilon)) = 0$ for all ball $\mathcal{B}(x,\epsilon)$ centered in $x \in A$ and of radius $\epsilon > 0$. Let $y \in (0,1)$ and $\mathfrak{D}_y = \{f : A \to [0,1] \mid f$ is measurable and $\int_A f(x)\mathrm{d}\mu = y\}$. Then a function H_1 defined by $H_1(\epsilon) = \int_{\mathcal{B}(x,\epsilon)\cap E_f^c} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,\epsilon)\cap E_f^c}(1-f(z))\mathrm{d}\mu$ is a continuous function where $E_f = \{x \in A | f(x) > y\}$ and $f \in \mathfrak{D}_y$.*

Proof. Let $y \in \mathfrak{D}_y$. Consider

$$\begin{aligned}
H_1(a) &= \int_{\mathcal{B}(x,a)\cap E_f^c} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,a)\cap E_f^c} (1-f(z))\mathrm{d}\mu \\
&= \int_{\mathcal{B}(x,a)\cap E_f^c} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,a)\cap E_f^c} 1\mathrm{d}\mu + \int_{\mathcal{B}^c(x,a)\cap E_f^c} f(z)\mathrm{d}\mu \\
&= \int_{\mathcal{B}(x,a)\cap E_f^c} f(z)\mathrm{d}\mu + \int_{\mathcal{B}^c(x,a)\cap E_f^c} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,a)\cap E_f^c} 1\mathrm{d}\mu \\
&= \int_{E_f^c} f(z)\mathrm{d}\mu - \mu(\mathcal{B}^c(x,a)\cap E_f^c) \\
&= \int_{E_f^c} f(z)\mathrm{d}\mu - \big(1-\mu(\mathcal{B}(x,a)) + \mu(E_f^c) - \mu(\mathcal{B}^c(x,a)\cup E_f^c)\big) \\
&= \int_{E_f^c} f(z)\mathrm{d}\mu - 1 + \mu(\mathcal{B}(x,a)) - \mu(E_f^c) + \mu(\mathcal{B}^c(x,a)\cup E_f^c) \\
&= \int_{E_f^c} f(z)\mathrm{d}\mu + \mu(\mathcal{B}(x,a)) - \mu(E_f^c) + \big(-1+\mu(\mathcal{B}^c(x,a)\cup E_f^c)\big) \\
&= \int_{E_f^c} f(z)\mathrm{d}\mu + \mu(\mathcal{B}(x,a)) - \mu(E_f^c) - \big(1-\mu(\mathcal{B}^c(x,a)\cup E_f^c)\big) \\
&= \int_{E_f^c} f(z)\mathrm{d}\mu + \mu(\mathcal{B}(x,a)) - \mu(E_f^c) - \mu\left(\mathcal{B}(x,a)\cap E_f\right).
\end{aligned}$$

Let $b_n \nearrow a$.
Then $\mathcal{B}(x,b_n) \subseteq \mathcal{B}(x,a)$ and $\bigcup_n \mathcal{B}(x,b_n) = \mathcal{B}(x,a)$.
So, $\mu(\mathcal{B}(x,b_n)) \to \mu(\mathcal{B}(x,a))$ and $\mu(\mathcal{B}(x,b_n)\cap E_f) \to \mu(\mathcal{B}(x,a)\cap E_f)$.
Thus, $\lim_{b_n\to a^-} H_1(b_n) = H_1(a)$.
If $b_n \searrow a$, then $\bigcap_n \mathcal{B}(x,b_n) = \overline{\mathcal{B}(x,a)}$ and $\mu\left(\overline{\mathcal{B}(x,a)} \setminus \mathcal{B}(x,a)\right) = 0$, we can conclude that $\mu(\mathcal{B}(x,b_n)) \searrow \mu(\mathcal{B}(x,a))$ and $\mu(\mathcal{B}(x,b_n)\cap E_f) \searrow \mu(\mathcal{B}(x,a)\cap E_f)$.
Thus, $\lim_{b_n\to a^+} H_1(b_n) = H_1(a)$.
Hence, H_1 is a continuous function. □

Lemma 3.1.2. *Let A be a metric space and μ be a Borel probability measure on A such that $\mu(\overline{\mathcal{B}(x,\epsilon)} \setminus \mathcal{B}(x,\epsilon)) = 0$ for all ball $\mathcal{B}(x,\epsilon)$ centered in $x \in A$ and of radius $\epsilon > 0$. Let $y \in (0,1)$ and $\mathfrak{D}_y = \{f : A \to [0,1] \mid f$ is measurable and $\int_A f(x)\mathrm{d}\mu = y\}$. Then a function H_2 defined by $H_2(\epsilon) = \int_{\mathcal{B}(x,\epsilon)\cap E_f} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,\epsilon)\cap E_f}(1-f(z))\mathrm{d}\mu$ is a continuous function where $E_f = \{x \in A | f(x) > y\}$ and $f \in \mathfrak{D}_y$.*

Proof. Let $y \in \mathfrak{D}_y$. Consider

$$\begin{aligned}
H_2(a) &= \int_{\mathcal{B}(x,a)\cap E_f} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,a)\cap E_f} (1-f(z))\mathrm{d}\mu \\
&= \int_{\mathcal{B}(x,a)\cap E_f} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,a)\cap E_f} 1\mathrm{d}\mu + \int_{\mathcal{B}^c(x,a)\cap E_f} f(z)\mathrm{d}\mu \\
&= \int_{\mathcal{B}(x,a)\cap E_f} f(z)\mathrm{d}\mu + \int_{\mathcal{B}^c(x,a)\cap E_f} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,a)\cap E_f} 1\mathrm{d}\mu \\
&= \int_{E_f} f(z)\mathrm{d}\mu - \mu(\mathcal{B}^c(x,a)\cap E_f) \\
&= \int_{E_f} f(z)\mathrm{d}\mu - (1 - \mu(\mathcal{B}(x,a)) + \mu(E_f) - \mu(\mathcal{B}^c(x,a)\cup E_f)) \\
&= \int_{E_f} f(z)\mathrm{d}\mu - 1 + \mu(\mathcal{B}(x,a)) - \mu(E_f) + \mu(\mathcal{B}^c(x,a)\cup E_f) \\
&= \int_{E_f} f(z)\mathrm{d}\mu + \mu(\mathcal{B}(x,a)) - \mu(E_f) + (-1 + \mu(\mathcal{B}^c(x,a)\cup E_f)) \\
&= \int_{E_f} f(z)\mathrm{d}\mu + \mu(\mathcal{B}(x,a)) - \mu(E_f) - (1 - \mu(\mathcal{B}^c(x,a)\cup E_f)) \\
&= \int_{E_f} f(z)\mathrm{d}\mu + \mu(\mathcal{B}(x,a)) - \mu(E_f) - \mu\left(\mathcal{B}(x,a)\cap E_f^c\right).
\end{aligned}$$

Let $b_n \nearrow a$.
Then $\mathcal{B}(x,b_n) \subseteq \mathcal{B}(x,a)$ and $\bigcup_n \mathcal{B}(x,b_n) = \mathcal{B}(x,a)$.
So, $\mu(\mathcal{B}(x,b_n)) \to \mu(\mathcal{B}(x,a))$ and $\mu(\mathcal{B}(x,b_n)\cap E_f^c) \to \mu(\mathcal{B}(x,a)\cap E_f^c)$.
Thus, $\lim_{b_n\to a^-} H_2(b_n) = H_2(a)$.
If $b_n \searrow a$, then $\bigcap_n \mathcal{B}(x,b_n) = \overline{\mathcal{B}(x,a)}$ and $\mu\left(\overline{\mathcal{B}(x,a)} \setminus \mathcal{B}(x,a)\right) = 0$, we can conclude that $\mu(\mathcal{B}(x,b_n)) \searrow \mu(\mathcal{B}(x,a))$ and $\mu(\mathcal{B}(x,b_n)\cap E_f^c) \searrow \mu(\mathcal{B}(x,a)\cap E_f^c)$.
Thus, $\lim_{b_n\to a^+} H_2(b_n) = H_2(a)$.
Hence, H_2 is a continuous function. □

Lemma 3.1.3. *Let A be a metric space and μ be a Borel probability measure on A such that $\mu(\overline{\mathcal{B}(x,\epsilon)} \setminus \mathcal{B}(x,\epsilon)) = 0$ for all ball $\mathcal{B}(x,\epsilon)$ centered in $x \in A$ and of radius $\epsilon > 0$.*
Let $y \in (0,1)$ and $\mathfrak{D}_y = \{f : A \to [0,1] \mid f$ is measurable and $\int_A f(x)\mathrm{d}\mu = y\}$.
The supremum of $\int_A |f(x) - y|\mathrm{d}\mu$ over $f \in \mathfrak{D}_y$ happens when $f \in \{0,1\}$ a.e.
Moreover, $\max\limits_{\{f\in\mathfrak{D}_y\}} \int_A |f(x) - y|\mathrm{d}\mu = 2y(1-y)$.

Proof. For each $y \in (0,1)$, we can find ball $\mathcal{B}(x_0,\epsilon_0) \subseteq A$ such that $\int_{\mathcal{B}(x_0,\epsilon_0)} \mathrm{d}\mu = y$ by continuity of the function $\epsilon \mapsto \mu(\mathcal{B}(x,\epsilon))$ with infimum zero and supremum one. We define a function $f : A \to \{0,1\}$ via

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathcal{B}(x_0,\epsilon_0), \\ 0 & \text{if } x \notin \mathcal{B}(x_0,\epsilon_0). \end{cases}$$

Since $\mu(\overline{\mathcal{B}(x,\epsilon)} \setminus \mathcal{B}(x,\epsilon)) = 0$, we can conclude that f is an indicator function in $\mathfrak{D}_y$.
Let $\mathbf{1}_B \in \mathfrak{D}_y$ be an indicator function of $B \subseteq A$. We shall show that

$$\int_A |\mathbf{1}_B(x) - y| \,\mathrm{d}\mu = 2y(1-y).$$

Since $\mathbf{1}_B \in \mathfrak{D}_y$, we can conclude that $\mu(B) = \int \mathbf{1}_B \mathrm{d}\mu = y$. Consider,

$$\begin{aligned} \int_A |\mathbf{1}_B(x) - y| \,\mathrm{d}\mu &= \int_B (1-y)\mathrm{d}\mu + \int_{B^c} y\mathrm{d}\mu \\ &= (1-y)\mu(B) + y\mu(B^c) \\ &= (1-y)y + y(1-y) \\ &= 2y(1-y). \end{aligned} \tag{3.1}$$

Therefore, $\int_A |\mathbf{1}_B(x) - y| \,\mathrm{d}\mu = 2y(1-y)$.
Let $f \in \mathfrak{D}_y$ be not an indicator function and $E_f = \{x \in A | f(x) > y\}$.
Since

$$\begin{aligned} \int_{E_f} y\mathrm{d}\mu + \int_{E_f^c} y\mathrm{d}\mu &= y \\ &= \int_A f(x)\mathrm{d}\mu \\ &= \int_{E_f} f(x)\mathrm{d}\mu + \int_{E_f^c} f(x)\mathrm{d}\mu, \end{aligned}$$

we get

$$\int_{E_f^c} (y - f(x))\mathrm{d}\mu = \int_{E_f} (f(x) - y)\mathrm{d}\mu. \tag{3.2}$$

Consider

$$\begin{aligned}\int_A |f(x)-y|\mathrm{d}\mu &= \int_{E_f} |f(x)-y|\mathrm{d}\mu + \int_{E_f^c} |f(x)-y|\mathrm{d}\mu \\ &= \int_{E_f} (f(x)-y)\mathrm{d}\mu + \int_{E_f^c} (y-f(x))\mathrm{d}\mu \\ &= 2\int_{E_f} (f(x)-y)\mathrm{d}\mu.\end{aligned}$$

Since $\int_{E_f}(f(x)-y)\mathrm{d}\mu = \int_{E_f^c}(y-f(x))\mathrm{d}\mu$ and $\int_A |f(x)-y|\mathrm{d}\mu = 2\int_{E_f}(f(x)-y)d\mu$, we can conclude that $\int_A |f(x)-y|\mathrm{d}\mu = 2\int_{E_f^c}(y-f(x))\mathrm{d}\mu$.

Next, we show that $\int_A |f(x)-y|\mathrm{d}\mu \leq \int_A |f^*(x)-y|\mathrm{d}\mu$ where $F^* \in \mathfrak{D}_y$ is an indicator function.

Case i) $\int_{E_f^c} f(x)\mathrm{d}\mu > 0$.

Let $x \in E_f$ be fixed.

Since the function H_1 in Lemma 3.1.1 is continuous and $y \geq f(x)$ for all $x \in E_f^c$, we get

$$\begin{aligned}0 &> y-1 \\ &\geq f(x)-1 \\ &= -(1-f(x)).\end{aligned}$$

Then $H_1(0) = -\int_{E_f^c}(1-f(z))\mathrm{d}\mu < 0$ and $H_1(\infty) = \int_{E_f^c} f(z)\mathrm{d}\mu > 0$.

Thus, there exists $\epsilon_0 \in (0,\infty)$ such that $H_1(\epsilon_0) = 0$.

Define a new function f^* by $f^* = f\mathbf{1}_{E_f} + \mathbf{1}_{E_f^c \cap \mathcal{B}^c(x,\epsilon_0)}$.

We show that $f^* \in \mathfrak{D}_y$.

Since $H_1(\epsilon_0) = 0$, we have

$$\begin{aligned}0 &= \int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c} (1-f(z))\mathrm{d}\mu \\ &= \int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c} 1\mathrm{d}\mu + \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c} f(z)\mathrm{d}\mu.\end{aligned}$$

Therefore,

$$\int_{E_f^c} f(z)\mathrm{d}\mu = \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c} 1\mathrm{d}\mu. \tag{3.3}$$

Then

$$\begin{aligned}\int_A f^*(x)\mathrm{d}\mu &= \int_A f\mathbf{1}_{E_f}(x)\mathrm{d}\mu + \int_A \mathbf{1}_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c}(x)\mathrm{d}\mu \\ &= \int_{E_f} f(x)\mathrm{d}\mu + \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c} \mathrm{d}\mu \\ &= \int_{E_f} f(x)\mathrm{d}\mu + \int_{E_f^c} f(z)\mathrm{d}\mu \\ &= \int_A f(x)\mathrm{d}\mu \\ &= y.\end{aligned}$$

Hence, $f^* \in \mathfrak{D}_y$.

Next, we show that $\int_A |f(x)-y|\mathrm{d}\mu < \int_A |f^*(x)-y|\mathrm{d}\mu$.

Since $y \in (0,1)$ and equation(3.3), we obtain that

$$\begin{aligned}\int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu + \int_{E_f^c} f(x)\mathrm{d}\mu &= \int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu + \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c} \mathrm{d}\mu \\ &> \int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu + \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu \\ &= \int_{E_f^c} y\mathrm{d}\mu.\end{aligned}$$

Therefore, $\int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu + \int_{E_f^c} f(x)\mathrm{d}\mu > \int_{E_f^c} y\mathrm{d}\mu$.

Thus,

$$\int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu > \int_{E_f^c} (y - f(x))\mathrm{d}\mu. \tag{3.4}$$

Consider

$$\begin{aligned}
\int_A |f^*(x) - y|\mathrm{d}\mu &= \int_{E_f} |f^*(x) - y|\mathrm{d}\mu + \int_{E_f^c} |f^*(x) - y|\mathrm{d}\mu \\
&= \int_{E_f} |f\mathbf{1}_{E_f}(x) + \mathbf{1}_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c}(x) - y|\mathrm{d}\mu \\
&\quad + \int_{E_f^c} |f\mathbf{1}_{E_f}(x) + \mathbf{1}_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c}(x) - y|\mathrm{d}\mu \\
&= \int_{E_f} |f(x) - y|\mathrm{d}\mu + \int_{E_f^c} |\mathbf{1}_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c}(x) - y|\mathrm{d}\mu \\
&= \int_{E_f} |f(x) - y|\mathrm{d}\mu + \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c} (1-y)\mathrm{d}\mu + \int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu \\
&= \int_{E_f} f(x)\mathrm{d}\mu - \int_{E_f} y\mathrm{d}\mu + \int_{E_f^c} f(x)\mathrm{d}\mu \\
&\quad - \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu + \int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu \\
&= \int_A f(x)\mathrm{d}\mu - \int_{E_f} y\mathrm{d}\mu - \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu + \int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu \\
&= y - \int_{E_f} y\mathrm{d}\mu - \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu + \int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu \\
&= \int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu + \int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu \\
&= 2\int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu.
\end{aligned}$$

By inequality (3.4) and equality (3.1), we can conclude that

$$\begin{aligned}
\int_A |f^*(x) - y|\mathrm{d}\mu &= 2\int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} y\mathrm{d}\mu \\
&> 2\int_{E_f^c} (y - f(x))\mathrm{d}\mu \\
&= \int_A |f(x) - y|\mathrm{d}\mu.
\end{aligned}$$

Hence, $\int_A |f^*(x) - y|\mathrm{d}\mu > \int_A |f(x) - y|\mathrm{d}\mu$.

Case ii) $\int_{E_f^c} f(x)\mathrm{d}\mu = 0$ and $\int_{E_f} (1 - f(x))\mathrm{d}\mu > 0$.
The function H_2 in Lemma 3.1.2 is continuous on $(0, \infty)$.

Consider

$$\begin{aligned}
H_2(0) &= \int_{\mathcal{B}(x,0)\cap E_f} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,0)\cap E_f} (1-f(z))\mathrm{d}\mu \\
&= 0 - \int_{\mathcal{B}^c(x,0)\cap E_f} (1-f(z))\mathrm{d}\mu \\
&= -\int_{E_f} (1-f(z))\mathrm{d}\mu \\
&= \int_{E_f} (f(z)-1)\mathrm{d}\mu \\
&< 0.
\end{aligned}$$

We next show that $H_2(\infty) > 0$. Consider

$$\begin{aligned}
H_2(\infty) &= \int_{\mathcal{B}(x,\infty)\cap E_f} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,\infty)\cap E_f} (1-f(z))\mathrm{d}\mu \\
&= \int_{\mathcal{B}(x,\infty)\cap E_f} f(z)\mathrm{d}\mu \\
&= \int_{E_f} f(z)\mathrm{d}\mu \\
&> 0.
\end{aligned}$$

Then there is $\epsilon_0 \in (0,1)$ such that $H_2(\epsilon_0) = 0$. Consider

$$\begin{aligned}
0 &= H_2(\epsilon_0) \\
&= \int_{\mathcal{B}(x,\epsilon_0)\cap E_f} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f} (1-f(z))\mathrm{d}\mu \\
&= \int_{\mathcal{B}(x,\epsilon_0)\cap E_f} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f} 1\mathrm{d}\mu + \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f} f(z)\mathrm{d}\mu \\
&= \int_{E_f} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f} 1\mathrm{d}\mu.
\end{aligned}$$

Hence,

$$\begin{aligned}
\int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f} 1\mathrm{d}\mu &= \int_{E_f} f(z)\mathrm{d}\mu \\
&= y. \qquad (3.5)
\end{aligned}$$

We define a new function $f^* = f\mathbf{1}_{E_f^c} + \mathbf{1}_{\mathcal{B}^c(x,\epsilon_0)\cap E_f}$. We show that $f^* \in \mathfrak{D}_y$.

Consider

$$\begin{aligned}\int_A f^*(x)\mathrm{d}\mu &= \int_A f\mathbf{1}_{E_f^c}\mathrm{d}\mu + \int_A \mathbf{1}_{\mathcal{B}^c(x,\epsilon_0)\cap E_f}\mathrm{d}\mu \\ &= \int_{E_f^c} f(x)d\mu + \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f} 1\mathrm{d}\mu \\ &= 0 + \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f} 1\mathrm{d}\mu \\ &= \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f} 1\mathrm{d}\mu \\ &= y.\end{aligned}$$

Thus, $f^* \in \mathfrak{D}_y$.

We next show that $\int_A |f(x) - y|\mathrm{d}\mu < \int_A |f^*(x) - y|\mathrm{d}\mu$. By inequality (3.2) and

$\int_{E_f^c} f(x)\mathrm{d}\mu = 0$, we obtain that

$$\begin{aligned}
\int_A |f^*(x) - y|\mathrm{d}\mu &= \int_A |f\mathbf{1}_{E_f^c}(x) + \mathbf{1}_{\mathcal{B}^c(x,\epsilon_0)\cap E_f}(x) - y|\mathrm{d}\mu \\
&= \int_{E_f} |\mathbf{1}_{\mathcal{B}^c(x,\epsilon_0)\cap E_f}(x) - y|\mathrm{d}\mu + \int_{E_f^c} |f(x) - y|\mathrm{d}\mu \\
&= \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f} (1-y)\,\mathrm{d}\mu + \int_{\mathcal{B}(x,\epsilon_0)\cap E_f} y\mathrm{d}\mu + \int_{E_f^c} (y - f(x))\,\mathrm{d}\mu \\
&= \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f} 1\mathrm{d}\mu - \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f} y\mathrm{d}\mu + \int_{\mathcal{B}(x,\epsilon_0)\cap E_f} y\mathrm{d}\mu \\
&\quad + \int_{E_f^c} y\mathrm{d}\mu - \int_{E_f^c} f(x)\mathrm{d}\mu \\
&= y - \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f} y\mathrm{d}\mu + \int_{\mathcal{B}(x,\epsilon_0)\cap E_f} y\mathrm{d}\mu + \int_{E_f^c} y\mathrm{d}\mu - 0 \\
&= \int_{E_f^c} y\mathrm{d}\mu + \int_{\mathcal{B}(x,\epsilon_0)\cap E_f} y\mathrm{d}\mu + \int_{\mathcal{B}(x,\epsilon_0)\cap E_f} y\mathrm{d}\mu + \int_{E_f^c} y\mathrm{d}\mu \\
&= 2\left(\int_{E_f^c} y\mathrm{d}\mu + \int_{\mathcal{B}(x,\epsilon_0)\cap E_f} y\mathrm{d}\mu\right) \\
&> 2\int_{E_f^c} y\mathrm{d}\mu \\
&= 2\int_{E_f^c} (y - f(x))\mathrm{d}\mu \\
&= \int_{E_f} (f(x) - y)\mathrm{d}\mu + \int_{E_f^c} (y - f(x))\mathrm{d}\mu \\
&= \int_{E_f} f(x)\mathrm{d}\mu - \int_{E_f} y\mathrm{d}\mu + \int_{E_f^c} y\mathrm{d}\mu - \int_{E_f^c} f(x)\mathrm{d}\mu \\
&= \int_{E_f} |f(x) - y|\mathrm{d}\mu + \int_{E_f^c} |f(x) - y|\mathrm{d}\mu \\
&= \int_A |f(x) - y|\mathrm{d}\mu.
\end{aligned}$$

Thus, $\int_A |f(x) - y|\mathrm{d}\mu < \int_A |f^*(x) - y|\mathrm{d}\mu$. Therefore, the supremum of $\int_A |f(x) - y|\mathrm{d}\mu$ over $f \in \mathfrak{D}_y$ happens when $f \in \{0,1\}$ a.e.

□

Lemma 3.1.4. *Let Y be a continuous random vector with dimension n. Then*

$$F_{Y|Y}(\vec{y}|\vec{x}) = \begin{cases} 1 & \textit{if} \quad x_i < y_i \quad ; i = 1, ..., n \\ \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. Let Y be a continuous random vector of dimension n, we have two cases to consider

Case i) $\vec{x} < \vec{y}$.

Then $V_{F_{Y,Y}}((x-h, x+h] \times (-\infty, y]) = V_{F_Y}((x-h, x+h])$ where h is sufficiently small. Therefore,

$$\begin{aligned} F_{Y|Y}(y|x) &= \lim_{h \searrow 0} \frac{V_{F_{Y,Y}}((x-h, x+h] \times (-\infty, y])}{V_{F_Y}((x-h, x+h])} \\ &= \lim_{h \searrow 0} \frac{V_{F_Y}((x-h, x+h])}{V_{F_Y}((x-h, x+h])} \\ &= 1. \end{aligned}$$

Case ii) $x \not< y$.

If $x_i > y_i$ for some $i = 1, ..., n$, then $\lim_{h \searrow 0} V_{F_{Y,Y}}((x-h, x+h] \times (-\infty, y]) = 0$.
Assume $x_i \le y_i$ for all $i = 1, ..., n$ and $x_j = y_j$ for some $j = 1, ..., n$.
Since Y is a continuous random vector, we get
$\lim_{h \searrow 0} V_{F_{Y,Y}}((x-h, x+h] \times (-\infty, y]) = 0$.

□

Recall the definition of φ from Definition 3.1.1,

$$\varphi(Y|X) = \int \int \left| F_{Y|X}(v|u) - F_Y(v) \right| \mathrm{d}F_X(u) \mathrm{d}F_Y(v).$$

Lemma 3.1.5. *Let Y be a continuous random vector. Then*

$$\varphi(Y|Y) = 2 \int (1 - F_Y(y)) F_Y(y) \mathrm{d}F_Y(y).$$

Proof.

$$\begin{aligned} \varphi(Y|Y) &= \int \int |F_{Y|Y}(y|x) - F_Y(y)| \mathrm{d}F_Y(x) \mathrm{d}F_Y(y) \\ &= \int \int_{\{x<y\}} (1 - F_Y(y)) \, \mathrm{d}F_Y(x) \mathrm{d}F_Y(y) + \int \int_{\{x \not< y\}} F_Y(y) \mathrm{d}F_Y(x) \mathrm{d}F_Y(y) \\ &= \int \left((1 - F_Y(y)) \int_{\{x<y\}} \mathrm{d}F_Y(x) \right) \mathrm{d}F_Y(y) + \int \left(F_Y(y) \int_{\{x \not< y\}} \mathrm{d}F_Y(x) \right) \mathrm{d}F_Y(y) \\ &= \int (1 - F_Y(y)) F_Y(y) \mathrm{d}F_Y(y) + \int (1 - F_Y(y)) F_Y(y) \mathrm{d}F_Y(y) \\ &= 2 \int (1 - F_Y(y)) F_Y(y) \mathrm{d}F_Y(y). \end{aligned}$$

□

Lemma 3.1.6. *Let $X = (X_1, ..., X_n)$ be a continuous random vector. Then $\mathbb{P}(X \in \overline{\mathcal{B}(x,\epsilon)} \setminus \mathcal{B}(x,\epsilon)) = 0$ for all $x \in \mathbb{R}^n$ and $\epsilon > 0$.*

Proof. Let $x = (x_1^0, ..., x_n^0)$. We set $A_i = (x_1, ..., x_i - \epsilon, ..., x_n)$ and $B_i = (x_1, ..., x_i + \epsilon, ..., x_n)$ where $x_j \in [x_j^0 - \epsilon, x_j^0 + \epsilon]$ for all $j \neq i$. Thus, $\overline{\mathcal{B}(x,\epsilon)} \setminus \mathcal{B}(x,\epsilon) = \bigcup\limits_{i=1}^{n} (A_i \cup B_i)$. Since a random vector X is continuous, we get $\mathbb{P}(A_i \cup B_i) \leq \mathbb{P}(\mathbb{R}^{j-1} \times \{x_i^j\} \times \mathbb{R}^{n-j}) = 0$. Therefore, $\mathbb{P}(X \in \overline{\mathcal{B}(x,\epsilon)} \setminus \mathcal{B}(x,\epsilon)) = 0$. □

Theorem 3.1.7. *Let X and Y be continuous random vectors. Then $0 \leq \varphi(Y|X) \leq \varphi(Y|Y)$.*

Proof. It is easy to see that $0 \leq \varphi(Y|X)$. Next, we show that $\varphi(Y|X) \leq \varphi(Y|Y)$. By Lemma 3.1.3, we obtain that

$$2\int (1 - F_Y(y))F_Y(y)\mathrm{d}F_Y(y) \geq \int\int |F_{Y|X}(y|x) - F_Y(y)|\mathrm{d}F_X(x)\mathrm{d}F_Y(y).$$

By Lemma 3.1.5, we have

$$\begin{aligned}\varphi(Y|Y) &= 2\int (1 - F_Y(y))F_Y(y)\mathrm{d}F_Y(y)\\ &\geq \int |F_{Y|X}(y|x) - F_Y(y)|\mathrm{d}F_X(x)\mathrm{d}F_Y(y)\\ &= \varphi(Y|X).\end{aligned}$$

□

Theorem 3.1.8. *Let X and Y be continuous random vectors. Then $\varphi(Y|X) = 0$ if and only if X and Y are independent.*

Proof. Assume that $\varphi(Y|X) = 0$. Then $\int\int \left|F_{Y|X}(v|u) - F_Y(v)\right| \mathrm{d}F_X(u)\mathrm{d}F_Y(v) = 0$. Thus, $F_{Y|X}(v|u) - F_Y(v) = 0$. Hence, $F_{Y|X}(v|u) = F_Y(v)$. Therefore, X and Y are independent.

Conversely, assume X and Y are independent. We get $F_{Y|X}(v|u) = F_Y(v)$. Thus,

$$\begin{aligned}\varphi(Y|X) &= \int \left|F_{Y|X}(v|u) - F_Y(v)\right| \mathrm{d}F_X(u)\mathrm{d}F_Y(v)\\ &= \int \left|F_Y(v) - F_Y(v)\right| \mathrm{d}F_X(u)\mathrm{d}F_Y(v)\\ &= 0.\end{aligned}$$

□

Theorem 3.1.9. *Let X and Y be continuous random vectors. Then $\varphi(Y|X) = \varphi(Y|Y)$ if and only if Y is a function of X.*

Proof. Assume that $\varphi(Y|X) = \varphi(Y|Y)$.
Then $\int\int |F_{Y|X}(y|x) - F_Y(y)|\mathrm{d}F_X(x)\mathrm{d}F_Y(y) = \int 2(1 - F_Y(y))F_Y(y)\mathrm{d}F_Y(y)$.
Since $\int |F_{Y|X}(y|x) - F_Y(y)|\mathrm{d}F_X \leq 2(1 - F_Y(y))F_Y(y)$, we can conclude that $\int |F_{Y|X}(y|x) - F_Y(y)|\mathrm{d}F_X = 2(1 - F_Y(y))F_Y(y)$.
By Lemma 3.1.3 and Lemma 2.2.4, we have $F_{Y|X} \in \{0, 1\}$ a.e.
Since $F_{Y|X}(\cdot|x)$ is a distribution function, we obtain that the set $\{y|F_{Y|X}(y|x) = 1\}$ is closed under the pointwise minimum and closed under Euclidean topology.
Therefore, $\{y|F_{Y|X}(y|x) = 1\}$ has a minimum, we say $f(x)$.
Then $F_{Y|X}(\cdot|x) = \mathbf{1}_{[f(x),\infty]}$. Thus, $\mathbb{P}(Y = f(X)) = 1$.
Hence, Y is a function of X. □

In fact, $\varphi(Y|X)$ is not generally less than one so, we normalize it into the form

$$\frac{\varphi(Y|X)}{\varphi(Y|Y)} = \frac{\int \left|F_{Y|X}(y|x) - F_Y(y)\right| \mathrm{d}F_X(x)\mathrm{d}F_Y(y)}{2\int F_Y(y)(1 - F_Y(y))\mathrm{d}F_Y(y)}.$$

For any random vectors X and Y with the joint distribution function $F_{X,Y}$, we denote $\varphi(F_{X,Y}|X) = \varphi(Y|X)$.

Let F and G be joint distribution functions with the same marginals, then the convex combination $tF + (1-t)G$ also has the same marginals as F and G.

Theorem 3.1.10. *Let F and G be joint distribution functions with the same marginals. Let $(X_1, Y_1), (X_2, Y_2)$ and (X_t, Y_t) have the distribution functions F, G and $tF + (1-t)G$, respectively. Then $\varphi(Y_t|X_t) \leq t\varphi(Y_1|X_1) + (1-t)\varphi(Y_2|X_2)$.*

Proof. Let $H = tF + (1-t)G$.
Then $H_{Y_t|X_t} = tF_{Y_1|X_1} + (1-t)G_{Y_2|X_2}$ and $H_{Y_t} = tF_{Y_1} + (1-t)G_{Y_2}$. Thus,

$$\begin{aligned}
\varphi(H|X_t) &= \int |H_{Y_t|X_t}(y|x) - H_{Y_t}(y)|\mathrm{d}F_X(x)\mathrm{d}F_Y(y) \\
&= \int |tF_{Y_1|X_1}(y|x) + (1-t)G_{Y_2|X_2}(y|x) - tF_{Y_1}(y) - (1-t)G_{Y_2}(y)|\mathrm{d}F_X(x)\mathrm{d}F_Y(y) \\
&= \int |t(F_{Y_1|X_1}(y|x) - F_{Y_1}(y)) + (1-t)(G_{Y_2|X_2}(y|x) - G_{Y_2}(y))|\mathrm{d}F_X(x)\mathrm{d}F_Y(y) \\
&\leq t\int |F_{Y_1|X_1}(y|x) - F_{Y_1}(y)|\mathrm{d}F_X(x)\mathrm{d}F_Y(y) \\
&\quad + (1-t)\int |G_{Y_2|X_2}(y|x) - G_{Y_2}(y)|\mathrm{d}F_X(x)\mathrm{d}F_Y(y) \\
&= t\varphi(F|X_1) + (1-t)\varphi(G|X_2).
\end{aligned}$$

$\square$

By Theorem 3.1.10, we get $\varphi(tF_{X,Y}+(1-t)G_{X,Y}|X) \leq \max\{\varphi(F_{X,Y}|X), \varphi(G_{X,Y}|X)\}$ for all joint distributions $F_{X,Y}$ and $G_{X,Y}$ with the same marginals F_X and F_Y.

Theorem 3.1.11. *Let X and Y be random vectors.*
Then $\varphi(Y|f(X)) \leq \varphi(Y|X)$ for all measurable functions f.

Proof. Let $f(X) = Z$. We get

$$\begin{aligned}
\varphi(Y|X) &= \int\int |F_{Y|X}(y|x) - F_Y(y)| \mathrm{d}F_X(x)\mathrm{d}F_Y(y) \\
&= \int\int\int |F_{Y|X}(y|x) - F_Y(y)| \mathrm{d}F_{X|Z}(x|z)\mathrm{d}F_Z(z)\mathrm{d}F_Y(y) \\
&\geq \int\int |\int F_{Y|X}(y|x)\mathrm{d}F_{X|Z}(x|z) - F_Y(y)| \mathrm{d}F_Z(z)\mathrm{d}F_Y(y) \\
&= \int\int |F_{Y|Z}(y|z) - F_Y(y)| \mathrm{d}F_Z(z)\mathrm{d}F_Y(y) \\
&= \varphi(Y|Z).
\end{aligned}$$

Hence, $\varphi(Y|f(X)) \leq \varphi(Y|X)$ for all measurable functions f. $\square$

Theorem 3.1.12. *Let X, Y and Z be random vectors. Then $\varphi(Y, Y, Z|X) = \varphi(Y, Z|X)$.*

Proof. For any random vectors X, Y and Z we have

$$\begin{aligned}
\varphi(Y, Y, Z|X) &= \int\int |F_{Y,Y,Z|X}(y, w, z|x) - F_{Y,Y,Z}(y, w, z)| \mathrm{d}F_{Y,Y,Z}(y, w, z)\mathrm{d}F_X(x) \\
&= \int\int |F_{Y,Z|X}(y, z|x) - F_{Y,Z}(y, z)| \mathrm{d}F_{Y,Z}(y, z)\mathrm{d}F_X(x) \\
&= \varphi(Y, Z|X).
\end{aligned}$$

Therefore, $\varphi(Y, Y, Z|X) = \varphi(Y, Z|X)$. $\square$

3.2 Measure of Complete Dependence Based on Linkages

Definition 3.2.1. Let X and Y be absolutely continuous random vectors of dimensions m and n, respectively and $p \in [1, \infty)$.
We define the measure ζ_p of complete dependence by

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

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

Theorem 3.2.1. *Let X and Y be two absolutely continuous random vectors and f be a measurable function such that $f(X)$ has an absolutely continuous distribution function. Then $\zeta_p(Y|f(X)) \leq \zeta_p(Y|X)$.*

Proof. Since $\zeta_p(Y|X) = \zeta_p(\Psi_{F_Y}(Y)|\Psi_{F_X}(X))$, we may as well assume that X and Y are uniformly distributed. Similarly, the fact that $\zeta_p(Y|f(X)) = \zeta_p(Y|\Psi_{F_{f(X)}}(f(X))$ allows us to assume that $Z = f(X)$ has uniform distribution also. Now,

$$\begin{aligned}\frac{\partial}{\partial \vec{z}} C_{Z,Y}(\vec{z}, \vec{y}) &= F_{Y|Z}(\vec{y}|\vec{z}) \\ &= \int F_{Y|X}(\vec{y}|\vec{x}) \mathrm{d}F_{X|Z}(\vec{x}|\vec{z}) \\ &= \int \frac{\partial}{\partial \vec{x}} C_{X,Y}(\vec{x}, \vec{y}) \mathrm{d}F_{X|Z}(\vec{x}|\vec{z}).\end{aligned}$$

By Jensen's inequality, we have

$$\begin{aligned}\int\int \left|\frac{\partial}{\partial \vec{z}} C_{Z,Y}(\vec{z}, \vec{y}) - \Pi(\vec{y})\right|^p \mathrm{d}\vec{z}\mathrm{d}\vec{y} &= \int\int \left|\int \frac{\partial}{\partial \vec{x}} C_{X,Y}(\vec{x}, \vec{y}) \mathrm{d}F_{X|Z}(\vec{x}|\vec{z}) - \Pi(\vec{y})\right|^p \mathrm{d}\vec{z}\mathrm{d}\vec{y} \\ &= \int\int \left|\int \left(\frac{\partial}{\partial \vec{x}} C_{X,Y}(\vec{x}, \vec{y}) - \Pi(\vec{y})\right) \mathrm{d}F_{X|Z}(\vec{x}|\vec{z})\right|^p \mathrm{d}\vec{z}\mathrm{d}\vec{y} \\ &\leq \int\int\int \left|\frac{\partial}{\partial \vec{x}} C_{X,Y}(\vec{x}, \vec{y}) - \Pi(\vec{y})\right|^p \mathrm{d}F_{X|Z}(\vec{x}|\vec{z})\mathrm{d}\vec{z}\mathrm{d}\vec{y} \\ &= \int\int \left|\frac{\partial}{\partial \vec{x}} C_{X,Y}(\vec{x}, \vec{y}) - \Pi(\vec{y})\right|^p \mathrm{d}F_X(\vec{x})\mathrm{d}\vec{y},\end{aligned}$$

that is, $\zeta_p(Y|Z) \leq \zeta_p(Y|X)$. □

Lemma 3.2.2. *Let X and Y be two absolutely continuous random vectors. Then the three following properties are equivalent:*

i) $\zeta_p(Y|X) = 0$,

ii) $\frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u}, \vec{v}) = \Pi(\vec{v})$, *and*

iii) $C_{X,Y}(\vec{u}, \vec{v}) = \Pi(\vec{u})\Pi(\vec{v})$.

Proof. *i)* $\Rightarrow$ *ii)*. Assume that $\zeta_p(Y|X) = 0$.
Then $\left[\int_{[0,1]^m} \int_{[0,1]^n} \left|\frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u}, \vec{v}) - \Pi(\vec{v})\right|^p \mathrm{d}\vec{u}\mathrm{d}\vec{v}\right]^{\frac{1}{p}} = 0$.
Thus, $\int_{[0,1]^m} \int_{[0,1]^n} \left|\frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u}, \vec{v}) - \Pi(\vec{v})\right|^p \mathrm{d}\vec{u}\mathrm{d}\vec{v} = 0$.
Hence, $\left|\frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u}, \vec{v}) - \Pi(\vec{v})\right| = 0$.
Therefore, $\frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u}, \vec{v}) = \Pi(\vec{v})$.

$ii) \Rightarrow iii)$. Assume that $\dfrac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u},\vec{v}) = \Pi(\vec{v})$. Thus,

$$\begin{aligned} C_{X,Y}(\vec{a},\vec{b}) &= \int\limits_{[\vec{0},\vec{a}]} \frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u},\vec{b}) \mathrm{d}\vec{u} \\ &= \int\limits_{[\vec{0},\vec{a}]} \Pi(\vec{b}) \mathrm{d}\vec{u} \\ &= \Pi(\vec{a})\Pi(\vec{b}). \end{aligned}$$

$iii) \Rightarrow i)$. Assume that $C_{X,Y}(\vec{u},\vec{v}) = \Pi(\vec{u})\Pi(\vec{v})$. Then

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

□

Lemma 3.2.3. *Let X and Y be two absolutely continuous random vectors. Then X and Y are independent if and only if $\Psi_{F_X}(X)$ and $\Psi_{F_Y}(Y)$ are independent.*

Proof. Since X and Y are absolutely continuous random vectors, we have Ψ_{F_X} and Ψ_{F_Y} are invertible functions. Thus,

$$\begin{aligned} \mathbb{P}(\Psi_{F_X}(X) \in A, \Psi_{F_Y}(Y) \in B) &= \mathbb{P}(X \in \Psi_{F_X}^{-1}(A), Y \in \Psi_{F_Y}^{-1}(B)) \\ &= \mathbb{P}(X \in \Psi_{F_X}^{-1}(A))\mathbb{P}(Y \in \Psi_{F_Y}^{-1}(B)) \\ &= \mathbb{P}(\Psi_{F_X}(X) \in A)\mathbb{P}(\Psi_{F_Y} Y \in (B)). \end{aligned}$$

Therefore, $\Psi_{F_X}(X)$ and $\Psi_{F_Y}(Y)$ are independent.

Conversely, suppose that $\Psi_{F_X}(X)$ and $\Psi_{F_Y}(Y)$ are independent. Consider

$$\begin{aligned} \mathbb{P}(X \in A, Y \in B) &= \mathbb{P}(X \in \Psi_{F_X}^{-1}(\Psi_{F_X}(A)), Y \in \Psi_{F_Y}^{-1}(\Psi_{F_Y}(B))) \\ &= \mathbb{P}(\Psi_{F_X}(X) \in \Psi_{F_X}(A), \Psi_{F_Y}(Y) \in \Psi_{F_Y}(B)) \\ &= \mathbb{P}(\Psi_{F_X}(X) \in \Psi_{F_X}(A))\mathbb{P}(\Psi_{F_Y}(Y) \in \Psi_{F_Y}(B)) \\ &= \mathbb{P}(X \in A)\mathbb{P}(Y \in B). \end{aligned}$$

Hence, X and Y are independent. □

Theorem 3.2.4. *Let X and Y be two absolutely continuous random vectors. Then $\zeta_p(Y|X) = 0$ if and only if X and Y are independent.*

Proof. Assume that $\zeta_p(Y|X) = 0$. By Lemma 3.2.2, we have $C_{X,Y}(\vec{u},\vec{v}) = \Pi(\vec{u})\Pi(\vec{v})$. Since $C_{X,Y}$ is the joint distribution of $\Psi_{F_X}(X)$ and $\Psi_{F_Y}(Y)$, we obtain that $\Psi_{F_X}(X)$ and $\Psi_{F_Y}(Y)$ are independent. By Lemma 3.2.3, we get X and Y are independent.

Conversely, let X and Y be independent. By Lemma 3.2.2, we get $\Psi_{F_X}(X)$ and $\Psi_{F_Y}(Y)$ are independent. Thus, $C_{X,Y} = \Pi(\vec{u})\Pi(\vec{v})$. By Lemma 3.2.3, we can conclude that $\zeta_p(Y|X) = 0$. $\square$

Lemma 3.2.5. *Let A be a metric space and μ be a Borel probability measure on A such that $\mu(\overline{\mathcal{B}(x,\epsilon)}\backslash\mathcal{B}(x,\epsilon)) = 0$ for all ball $\mathcal{B}(x,\epsilon)$ with centered in $x \in A$ and of radius $\epsilon > 0$. Let $y \in (0,1)$ and $\mathfrak{D}_y = \{f : A \to [0,1] \mid f$ is measurable and $\int_A f(x)\mathrm{d}\mu = y\}$. Then f maximizes the function $\hat{f} \mapsto \int \left|\hat{f}(x) - y\right|^p \mathrm{d}\mu$ on $\mathfrak{D}_y$ if and only if f is an indicator function. Moreover, $\max\limits_{\{f\in\mathfrak{D}_y\}} \int_A |f(x) - y|^p \mathrm{d}\mu = y^p(1-y) + y(1-y)^p$.*

Proof. By symmetry, we may assume $0 < y \leq \frac{1}{2}$. By following the proof of Lemma3.1.3, we have $\mathfrak{D}_y$ contains an indicator function.
Let $B \subseteq A$ and $\mathbf{1}_B \in \mathfrak{D}_y$ be an indicator function of B. We shall show that

$$\int_A |\mathbf{1}_B(x) - y|^p \,\mathrm{d}\mu = (1-y)^p y + y^p(1-y).$$

Since $\mathbf{1}_B \in \mathfrak{D}_y$, we can conclude that $\mu(B) = \int \mathbf{1}_B \mathrm{d}\mu = y$. Consider

$$\begin{aligned}\int_A |\mathbf{1}_B(x) - y|^p \,\mathrm{d}\mu &= \int_B (1-y)^p \mathrm{d}\mu + \int_{B^c} y^p \mathrm{d}\mu \\ &= (1-y)^p \mu(B) + y^p \mu(B^c) \\ &= (1-y)^p y + y^p(1-y).\end{aligned}$$

Therefore, $\int_A |\mathbf{1}_B(x) - y|^p \,\mathrm{d}\mu = (1-y)^p y + y^p(1-y)$.
Let $f \in \mathfrak{D}_y$ and $E_f = \{x \in A | f(x) > y\}$.
Assume that f is not an indicator function.
We will construct another function $f^* \in \mathfrak{D}_y$ such that $\int_A |f(x)-y|^p \mathrm{d}\mu < \int_A |f^*(x)-y|^p \mathrm{d}\mu$.
Hence, f can not maximize $\hat{f} \mapsto \int |\hat{f}(x) - y|^p \mathrm{d}\mu$ on $\mathfrak{D}_y$.
Since $f \in \mathfrak{D}_y$, $\int_{E_f^c} (y - f(x))\mathrm{d}\mu = \int_{E_f} (f(x) - y)\mathrm{d}\mu$.
If $\int_{E_f^c} f(x)\mathrm{d}\mu = 0$, then $\int_{E_f} f(x)\mathrm{d}\mu = y$ and hence, $f = y$ a.e. which immediately implies f is not the maximizer.
Thus, we may assume $\int_{E_f^c} f(x)\mathrm{d}\mu > 0$.

For any $\epsilon > 0$, define $H(\epsilon) = \int_{\mathcal{B}(x,\epsilon)\cap E_f^c} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,\epsilon)\cap E_f^c}(1 - f(z))\mathrm{d}\mu$.
Since $0 > y - 1 \geq -(1 - f(x))$, we get $H(0) = -\int_{E_f^c}(1 - f(z))\mathrm{d}\mu < 0$.
Clearly, $\lim_{\epsilon\to\infty} H(\epsilon) = \int_{E_f^c} f(z)\mathrm{d}\mu > 0$.
By the assumption of μ, the function H is continuous.
Thus, there exists $\epsilon_0 \in (0, \infty)$ such that $H(\epsilon_0) = 0$.
Define a new function f^* by letting $f^* = f\mathbf{1}_{E_f} + \mathbf{1}_{E_f^c\cap\mathcal{B}^c(x,\epsilon_0)}$.

First, we show that $f^* \in \mathfrak{D}_y$. Since $H(\epsilon_0) = 0$, we have

$$\begin{aligned}
0 &= \int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c}(1 - f(z))\mathrm{d}\mu \\
&= \int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} f(z)\mathrm{d}\mu - \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c} 1\mathrm{d}\mu + \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c} f(z)\mathrm{d}\mu.
\end{aligned}$$

Therefore,

$$\int_{E_f^c} f(z)\mathrm{d}\mu = \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c} 1\mathrm{d}\mu. \tag{3.6}$$

Then

$$\begin{aligned}
\int_A f^*(x)\mathrm{d}\mu &= \int_A f\mathbf{1}_{E_f}(x)\mathrm{d}\mu + \int_A \mathbf{1}_{E_f^c\cap\mathcal{B}^c(x,\epsilon_0)}(x)\mathrm{d}\mu \\
&= \int_{E_f} f(x)\mathrm{d}\mu + \int_{E_f^c\cap\mathcal{B}^c(x,\epsilon_0)} 1\mathrm{d}\mu \\
&= \int_{E_f} f(x)\mathrm{d}\mu + \int_{E_f^c} f(z)\mathrm{d}\mu \\
&= \int_A f(x)\mathrm{d}\mu \\
&= y.
\end{aligned}$$

Hence, $f^* \in \mathfrak{D}_y$. Next, we show that $\int_A |f(x) - y|^p\mathrm{d}\mu < \int_A |f^*(x) - y|^p\mathrm{d}\mu$. Consider

$$\begin{aligned}
\int_A |f^*(x) - y|^p\mathrm{d}\mu &= \int_{E_f} |f^*(x) - y|^p\mathrm{d}\mu + \int_{E_f^c} |f^*(x) - y|^p\mathrm{d}\mu \\
&= \int_{E_f} |f\mathbf{1}_{E_f}(x) + \mathbf{1}_{E_f^c\cap\mathcal{B}^c(x,\epsilon_0)}(x) - y|^p\mathrm{d}\mu \\
&\quad + \int_{E_f^c} |f\mathbf{1}_{E_f}(x) + \mathbf{1}_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c}(x) - y|^p\mathrm{d}\mu \\
&= \int_{E_f} |f(x) - y|^p\mathrm{d}\mu + \int_{E_f^c} |\mathbf{1}_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c}(x) - y|^p\mathrm{d}\mu \\
&= \int_{E_f} |f(x) - y|^p\mathrm{d}\mu + \int_{\mathcal{B}^c(x,\epsilon_0)\cap E_f^c}(1 - y)^p\mathrm{d}\mu + \int_{\mathcal{B}(x,\epsilon_0)\cap E_f^c} y^p\mathrm{d}\mu.
\end{aligned}$$

Clearly, $|f(x) - y| \leq y \leq 1 - y$ for all $x \in E^c$. Thus,

$$\int_{E_f^c} |f(x) - y|^p\mathrm{d}\mu \leq \int_{B^c(x,\epsilon_0)\cap E_f^c}(1 - y)^p\mathrm{d}\mu + \int_{B(x,\epsilon_0)\cap E_f^c} y^p\mathrm{d}\mu$$

and both sides are equal only when $y = \dfrac{1}{2}$ and in this case $|f(x) - y| = y$ on E_f^c, that is, $f = 0$ on E_f^c. This contradicts the fact $\int_{E_f^c} f(x)\mathrm{d}\mu > 0$. Thus,

$$\int_{E_f^c} |f(x) - y|^p \mathrm{d}\mu < \int_{B^c(x,\epsilon_0)\cap E_f^c} (1-y)^p \mathrm{d}\mu + \int_{B(x,\epsilon_0)\cap E_f^c} y^p \mathrm{d}\mu.$$

Finally, we show that

$$\int_A |f(x) - y|^p \mathrm{d}\mu = (1-y)^p y + y^p(1-y)$$

whenever $f \in \mathfrak{D}_y$ is an indicator function. Consider

$$\begin{aligned}
\int_A |f(x) - y|^p \mathrm{d}\mu &= \int_{E_f} |f(x) - y|^p \mathrm{d}\mu + \int_{E_f^c} |f(x) - y|^p \mathrm{d}\mu \\
&= \int_{E_f} (1-y)^p \mathrm{d}\mu + \int_{E_f^c} y^p \mathrm{d}\mu \\
&= (1-y)^p \mu(E_f) + y^p \mu(E_f^c) \\
&= (1-y)^p y + y^p(1-y).
\end{aligned}$$

□

Corollary 3.2.6. *For any absolutely continuous random vector Y,*

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

Proof. Since $\dfrac{\partial}{\partial \vec{u}} C_{Y,Y}(\vec{u},\vec{v}) = C_{Y|Y}(\vec{v}|\vec{u}) = \mathbf{1}_{\{(\vec{u},\vec{v})|\vec{v}\leq\vec{u}\}}$ is an indicator function,

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

by Lemma 3.2.5. Thus,

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

□

Lemma 3.2.7. *Let Y be an absolutely continuous $m-$dimensional random vector and p be a positive integer. Then*

$$\zeta_p(Y|Y)^p = \sum_{k=0}^{p} \binom{p}{k} (-1)^k \frac{1}{(k+2)^m} + \frac{1}{(p+1)^m} - \frac{1}{(p+2)^m}.$$

Proof. Consider

$$\begin{aligned}
\zeta_p(Y|Y)^p &= \int_0^1 ... \int_0^1 \left(\Pi(y)(1-\Pi(y))^p + \Pi(y)^p(1-\Pi(y))\right) \mathrm{d}y_1...\mathrm{d}y_m \\
&= \int_0^1 ... \int_0^1 \left(\Pi(y)\sum_{k=0}^{p}\binom{p}{k}(-\Pi(y))^k + \Pi(y)^p - \Pi(y)^{p+1}\right) \mathrm{d}y_1...\mathrm{d}y_m \\
&= \int_0^1 ... \int_0^1 \left(\sum_{k=0}^{p}\binom{p}{k}(-1)^k\Pi(y)^{k+1} + \Pi(y)^p - \Pi(y)^{p+1}\right) \mathrm{d}y_1...\mathrm{d}y_m \\
&= \sum_{k=0}^{p}\binom{p}{k}(-1)^k\frac{1}{(k+2)^m} + \frac{1}{(p+1)^m} - \frac{1}{(p+2)^m}.
\end{aligned}$$

□

Remark 3.2.8. *Let Y be a random vector with dimension m, by Lemma 3.2.7, we get*

$$\begin{aligned}
\zeta_1(Y|Y) &= \sum_{k=0}^{1}\binom{1}{k}(-1)^k\frac{1}{(k+2)^m} + \frac{1}{(1+1)^m} - \frac{1}{(1+2)^m} \\
&= \binom{1}{0}\frac{1}{2^m} - \binom{1}{1}\frac{1}{3^m} + \frac{1}{2^m} - \frac{1}{3^m} \\
&= \frac{1}{2^m} - \frac{1}{3^m} + \frac{1}{2^m} - \frac{1}{3^m} \\
&= \frac{2}{2^m} - \frac{2}{3^m}
\end{aligned}$$

and

$$\begin{aligned}
\zeta_2^2(Y|Y) &= \sum_{k=0}^{2}\binom{2}{k}(-1)^k\frac{1}{(k+2)^m} + \frac{1}{(2+1)^m} - \frac{1}{(2+2)^m} \\
&= \binom{2}{0}\frac{1}{2^m} - \binom{2}{1}\frac{1}{3^m} + \binom{2}{2}\frac{1}{4^m} + \frac{1}{3^m} - \frac{1}{4^m} \\
&= \frac{1}{2^m} - \frac{2}{3^m} + \frac{1}{4^m} + \frac{1}{3^m} - \frac{1}{4^m} \\
&= \frac{1}{2^m} - \frac{1}{3^m}.
\end{aligned}$$

Therefore, $\zeta_1(Y|Y) = \frac{2}{2^m} - \frac{2}{3^m}$ *and* $\zeta_2(Y|Y) = \sqrt{\frac{2}{2^m} - \frac{2}{3^m}}$.

Theorem 3.2.9. *Let X and Y be two absolutely continuous random vectors. Then $0 \leq \zeta_p(Y|X) \leq \zeta_p(Y|Y)$.*

Proof. Clearly, $0 \leq \zeta_p(Y|X)$. Next, we show that $\zeta_p(Y|X) \leq \zeta_p(Y|Y)$.
By Lemma 3.2.5,

$$\int \left|\frac{\partial}{\partial \vec{u}}C_{X,Y}(\vec{u},\vec{v}) - \Pi(\vec{v})\right|^p \mathrm{d}\vec{u} \leq (\Pi(\vec{v}))^p(1-\Pi(\vec{v})) + \Pi(\vec{v})(1-\Pi(\vec{v}))^p.$$

Therefore,

$$
\begin{aligned}
\zeta_p(Y|X) &= \left(\int \int \left| \frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u}, \vec{v}) - \Pi(\vec{v}) \right|^p \mathrm{d}\vec{u}\mathrm{d}\vec{v} \right)^{\frac{1}{p}} \\
&\leq \left(\int [(\Pi(\vec{v}))^p (1 - \Pi(\vec{v})) + \Pi(\vec{v})(1 - \Pi(\vec{v}))^p] \, \mathrm{d}\vec{v} \right)^{\frac{1}{p}} \\
&= \zeta_p(Y|Y).
\end{aligned}
$$

□

Theorem 3.2.10. *Let X and Y be two absolutely continuous random vectors. The three following properties are equivalent:*

i) *Y is a measurable function of X,*

ii) *$\Psi_{F_Y}(Y)$ is a measurable function of $\Psi_{F_X}(X)$, and*

iii) *$\zeta_p(Y|X) = \zeta_p(Y|Y)$.*

Proof. First, we show that *i*) and *ii*) are equivalent. Assume Y is a measurable function of X, that is, $Y = f(X)$ for some measurable function f.
Thus,

$$\Psi_{F_Y}(Y) = \Psi_{F_{f(X)}}(f(X)) = \Psi_{F_{f(X)}} f(\Psi^{-1}_{F_{f(X)}}(\Psi_{F_{f(X)}}(X)),$$

that is, $\Psi_{F_{Y(Y)}}$ is a measurable function of $\Psi_{F_X}(X)$.
Conversely, assume $\Psi_{F_Y}(Y)$ is a measurable function of $\Psi_{F_X}(X)$.
Then $\Psi_{F_Y}(Y) = g(\Psi_{F_X}(X))$ for some measurable function g. So, $Y = \Psi^{-1}_{F_Y}(g(\Psi_{F_X}(X)))$.

Finally, we show that *i*) and *iii*) are equivalent.
Assume Y is a measurable function of X, that is, $Y = f(X)$ for some measurable function f. The fact that $\zeta_p(Y|X) = \zeta_p(Y|Y)$ immediately follows from Theorem 3.2.1 and Theorem 3.2.9.

Conversely, assume $\zeta_p(Y|X) = \zeta_p(Y|Y)$.
Then $\int \left| \frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u}, \vec{v}) - \Pi(\vec{v}) \right|^p \mathrm{d}\vec{u}\mathrm{d}\vec{v} = \int [(\Pi(\vec{v}))^p (1 - \Pi(\vec{v})) + \Pi(\vec{v})(1 - \Pi(\vec{v}))^p] \mathrm{d}\vec{v}$.
Thus, $\int \left| \frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u}, \vec{v}) - \Pi(\vec{v}) \right|^p \mathrm{d}\vec{u} = (\Pi(\vec{v}))^p (1 - \Pi(\vec{v})) + \Pi(\vec{v})(1 - \Pi(\vec{v}))^p$.
By Lemma 3.2.5, $\frac{\partial}{\partial \vec{u}} C_{X,Y}$ is an indicator function.
Since $\frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u}, \cdot) = C_{Y|X}(\cdot|\vec{u})$ is a distribution function, the set $\left\{ \vec{v} | \frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u}, \vec{v}) = 1 \right\}$ is closed under the pointwise minimum and closed under Euclidean topology.
Therefore, $\left\{ \vec{v} | \frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u}, \vec{v}) = 1 \right\}$ has a minimum, say $f(\vec{u})$.

Then $\dfrac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u},\vec{v}) = \mathbf{1}_{[f(\vec{u}),\infty)}(\vec{v})$.
Thus, $\mathbb{P}(Y = f(X) = 1)$, that is, Y is a measurable function of X. $\square$

Theorem 3.2.11. *Let X, Y, and Z be absolutely continuous random vectors in which Z has the dimension k and (X,Y) and Z are independent,*

$$\zeta_p(Y,Z|X) = \left(\frac{1}{p+1}\right)^{\frac{k}{p}} \zeta_p(Y|X).$$

In particular, $\zeta_p(Y,Z|X) < \zeta_p(Y|X)$.

Proof. Since (X,Y) and Z are independent, we have $C_{X,(Y,Z)}(\vec{u},(\vec{v},\vec{w})) = C_{X,Y}(\vec{u},\vec{v})\Pi(\vec{w})$. Thus,

$$\begin{aligned}
(\zeta_p(Y,Z|X))^p &= \int\int\int \left| \frac{\partial}{\partial \vec{u}} C_{X,(Y,Z)}(\vec{u},(\vec{v},\vec{w})) - \Pi(\vec{v})\Pi(\vec{w}) \right|^p \mathrm{d}\vec{u}\mathrm{d}\vec{v}\mathrm{d}\vec{w} \\
&= \int\int\int \left| \frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u},\vec{v})\Pi(\vec{w}) - \Pi(\vec{v})\Pi(\vec{w}) \right|^p \mathrm{d}\vec{u}\mathrm{d}\vec{v}\mathrm{d}\vec{w} \\
&= \int (\Pi(\vec{w}))^p \mathrm{d}\vec{w} \int\int \left| \frac{\partial}{\partial \vec{u}} C_{X,Y}(\vec{u},\vec{v}) - \Pi(\vec{v}) \right|^p \mathrm{d}\vec{u}\mathrm{d}\vec{v} \\
&= \left(\frac{1}{p+1}\right)^k \zeta_p(Y|X)^p.
\end{aligned}$$

Therefore, $\zeta_p(Y,Z|X) = \left(\frac{1}{p+1}\right)^{\frac{k}{p}} \zeta_p(Y|X)$. $\square$

Corollary 3.2.12. *Let X, Y and Z be absolutely continuous random vectors such that (X,Y) and Z are independent. Then $\dfrac{\zeta_k(Y,Z|X)}{\zeta_k(Y,Z|Y,Z)} < \dfrac{\zeta_k(Y|X)}{\zeta_k(Y|Y)}$.*

Proof. Assume that Y has the dimension m and Z has the dimension n. Consider

$$\begin{aligned}
\zeta_p^k(Y,Z|Y,Z) &= \int_{[0,1]^m}\int_{[0,1]^n} \Big[\Pi(\vec{u},\vec{v})(1-\Pi(\vec{u},\vec{v}))^k + \Pi^k(\vec{u},\vec{v})(1-\Pi(\vec{u},\vec{v}))\Big] \mathrm{d}\vec{v}\mathrm{d}\vec{u} \\
&= \int_{[0,1]^m}\int_{[0,1]^n} \Big[\Pi(\vec{u})\Pi(\vec{v})(1-\Pi(\vec{u})\Pi(\vec{v}))^k + \Pi^k(\vec{u})\Pi^k(\vec{v})(1-\Pi(\vec{u})\Pi(\vec{v})\Big] \mathrm{d}\vec{v}\mathrm{d}\vec{u} \\
&> \int_{[0,1]^m}\int_{[0,1]^n} \Big[\Pi(\vec{u})\Pi^k(\vec{v})(1-\Pi(\vec{u}))^k + \Pi^k(\vec{u})\Pi^k(\vec{v})(1-\Pi(\vec{u}))\Big] \mathrm{d}\vec{v}\mathrm{d}\vec{u} \\
&= \frac{1}{(p+1)^n}\int_{[0,1]^m} \Big[\Pi(\vec{u})(1-\Pi(\vec{u}))^k + \Pi^k(\vec{u})(1-\Pi(\vec{u}))\Big] \mathrm{d}\vec{u} \\
&= \frac{1}{(p+1)^n}\zeta_p^k(Y|Y).
\end{aligned}$$

Thus, $\zeta_p(Y,Z|Y,Z) > \frac{1}{(p+1)^{\frac{n}{p}}}\zeta_p(Y|Y)$.
By Theorem 3.2.11, we can conclude that $\dfrac{\zeta_p(Y,Z|X)}{\zeta_p(Y,Z|Y,Z)} < \dfrac{\zeta_p(Y|X)}{\zeta_p(Y|Y)}$. $\square$

Finally, we prove the next theorem which is a generalization of Theorem 1 in [7].

Theorem 3.2.13. *For any $\epsilon > 0$, there are absolutely continuous random vectors X and Y of arbitrary marginals but with the same dimension such that Y is completely dependent on X but $\zeta_p(X|Y) \leq \epsilon$.*

Proof. Since ζ_p only depends on linkages, it is sufficient to prove the result for a uniform distribution.

First, we prove the result in the case of random variables. Let $Y = kX \mod 1$ where X has a uniform distribution. Then Y is also uniformly distributed. Clearly, $\zeta_p(Y|X) = \zeta_p(Y|Y)$ since Y is a function of X. Also,

$$\begin{aligned}\frac{\partial}{\partial u} C_{Y,X}(u,v) &= \mathbb{P}(X \leq v | Y = u) \\ &= \begin{cases} \frac{1}{k}\lfloor kv - u \rfloor & \text{if } kv > u, \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

where $\lfloor x \rfloor$ is the floor of x. Thus, $\zeta_p^p(X|Y) \leq \frac{1}{k}\int |u|^p du$ and hence,

$$\begin{aligned}\zeta_p(X|Y) &\leq \left(\frac{1}{k}\right)^{\frac{1}{p}} \left(\frac{1}{p+1}\right)^{\frac{1}{p}} \\ &\leq \left(\frac{1}{k}\right)^{\frac{1}{p}}.\end{aligned}$$

Choosing $k = \lfloor \epsilon^{-p} \rfloor + 1$, we have $\zeta_p(X|Y) \leq \epsilon$ as required.

In general, let (X_i, Y_i) where $i = 1, ..., n$ be independent copies of a random vector (X_0, Y_0) in which $\zeta_p(Y_0|X_0) = \zeta_p(Y_0|Y_0)$ but $\zeta_p(X_0|Y_0) \leq \frac{\epsilon}{n}$ and let $X = (X_1, ..., X_n)$ and $Y = (Y_1, ..., Y_n)$. Then $\zeta_p(Y|X) = \zeta_p(Y|Y)$ and

$$\begin{aligned}\zeta_p^p(X|Y) &= \int \left| \Pi_{i=1}^n \frac{\partial}{\partial u_i} C_{Y_0,X_0}(u_i, v_i) - \Pi_{i=1}^n v_i \right|^p \mathrm{d}u_1 \cdots \mathrm{d}u_n \mathrm{d}v_1 \cdots \mathrm{d}v_n \\ &\leq \sum_{i=1}^n n^{p-1} \int \left| \frac{\partial}{\partial u} C_{Y_0,X_0}(u,v) - v \right|^p \mathrm{d}u\mathrm{d}v \\ &= n^p \zeta_p^p(X_0, |Y_0) \\ &= \epsilon^p\end{aligned}$$

and hence, $\zeta_p(X|Y) \leq \epsilon$ as desired. □