UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y
DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

SIMPLE TESTS OF MULTIVARIATE INDEPENDENCE USING THE SAMPLE COPULA

QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS

PRESENTA:
JUAN DIEGO NIEVES LEDESMA

DIRECTOR DE LA TESIS
Dr. JOSÉ MARÍA GONZÁLEZ-BARRIOS MURGUÍA
INSTITUTO DE INVESTIGACIONES EN MATEMÁTICAS
APLICADAS Y EN SISTEMAS

MÉXICO, CIUDAD DE MÉXICO, JUNIO 2018

Contents

Introduction

In this thesis, we present a characterization of independence using the d-sample copula defined in [9] and improved in [11] and using the checkerboard approximation. Using this characterization, we define some statistics based on different metrics such as the total variation distance, the Hellinger distance and the supremum distance; and we also define a statistic based on the divergence of Kullback-Leibler.

This work is divided in five chapters. In the first chapter we present the basic definitions and the most important results in copula theory.

In the second chapter we give a comparison between the empirical copula and the d-sample copula of order m. Moreover, we note that it is a better idea to use the sample copula instead the empirical copula, because of the simplicity of its evaluation.

In the third chapter we find the distribution of the sample copula under the assumption of independence and we establish a way to evaluate moments.

In the fourth chapter we present the most important result of the present work: we give a simple characterization of independence through the checkerboard approximation and define four statistics. We also compare our tests with some of the most used tests via simulations. Moreover, we use the proposed tests with real data. Finally, in the last section we propose further investigations.

In the fifth chapter we give our final conclusions.

Chapter 1

Preliminaries

We begin this chapter the basic definitions and main results of copulas. Moreover, in the second section, we will see a certain type of dependence, called concordance. Also, in the third section, we introduce an important class of copulas known as Archimedean. The results presented are based on Nelsen's book, *An Introduction to Copulas* [18]. We omit the proofs because they are well known results.

1.1 Basic definitions and main results

This section includes the basic definitions and the most important theorem in copula theory: Sklar's Theorem. We also present the definitions and the principal results of the Frechet-Hoeffding bounds and the product copula.

Definition 1.1.1. *Let* $\mathbf{I} = [0,1]$ *be the closed unit interval and let* $d \geq 2$ *be the dimension. Let* $S_1, S_2, ..., S_d$ *be subsets of* $\mathbf{I}$ *such that* $0, 1 \in S_i$ *for every* $i \in I_d$*, where* $I_d = \{1, 2, ..., d\}$*. Let* $C' : S_1 \times S_2 \times \cdots \times S_d \to \mathbb{R}$ *be a function. Then* C' *is a* **d-*subcopula*** *if and only if* C' *satisfies:*

i) $C'(u_1, ..., u_d) = 0$ *if at least one* $u_i = 0$ *for some* $i \in I_d$*;*

ii) $C'(1, ..., u_i, 1, ..., 1) = u_i$ *for every* $i \in I_d$ *and for every* $u_i \in S_i$*;*

iii) C' *is* **d-*increasing***, *that is, for every* $0 \leq u_i \leq v_i \leq 1$ *such that* $u_i, v_i \in S_i$ *for every* $i \in I_d$*, we have that if* $B = [u_1, v_1] \times \cdots \times [u_d, v_d]$ *then*

$$V_{C'}(B) := \sum_{\underline{b}} sgn(\underline{b})C'(\underline{b}) \geq 0, \tag{1.1}$$

where the sum runs over all $\underline{b} = (b_1, ..., b_d)$ *which are the vertices of* B*, and the sign function is defined by*

$$sgn(\underline{b}) = \begin{cases} 1 & \textit{if } b_k = u_k \textit{ for an even number of } k\textit{'s;} \\ -1 & \textit{if } b_k = u_k \textit{ for an odd number of } k\textit{'s.} \end{cases}$$

We say that a d-subcopula is a **d-*copula*** *if and only if* $S_1 = S_2 = \cdots = S_d = \mathbf{I}$.

The following theorem says that copulas have a Lipschitz condition on $\mathbf{I}^d$. From this it follows immediately that copulas are uniformly continuous.

Theorem 1.1.2. *Let* $S_1, S_2, ..., S_d$ *be subsets of* $\mathbf{I}$ *such that* $0, 1 \in S_i$ *for every* $i \in I_d$. *Let* $C' : S_1 \times S_2 \times \cdots \times S_d \to \mathbb{R}$ *be a subcopula. Then for every* $(u_1, ..., u_d), (v_1, ..., v_d) \in \mathbf{I}^d$,

$$|C'(u_1, ..., u_d) - C'(v_1, ..., v_d)| \leq \sum_{i=1}^{d} |v_i - u_i|,$$

i.e., C' *is uniformly continuous in* S.

The following theorem is the most important in the theory of copulas and was introduced by Sklar in his doctoral thesis in 1959.

Theorem 1.1.3. *(Sklar's Theorem) Let* H *be a joint d-distribution function for* $d \geq 2$ *with margins* $F_1, F_2, ..., F_d$. *Then there exists a d-copula* C *such that for every* $(x_1, x_2, ..., x_d) \in \mathbb{R}^d$,

$$H(x_1, x_2, ..., x_d) = C(F_1(x_1), F_2(x_2), ..., F_d(x_d)). \tag{1.2}$$

If $F_1, F_2, ..., F_d$ *are continuous, then* C *is unique; otherwise,* C *is uniquely determined on* $Ran(F_1) \times Ran(F_2) \times \cdots \times Ran(F_d)$. *Conversely, if* C *is a d-copula and* $F_1, F_2, ..., F_d$ *are distribution functions, then the function* H *defined in equation* (1.2) *is a joint d-distribution function in* $\mathbb{R}^d$.

Theorem 1.1.4. *Let* $(u_1, ..., u_d) \in \mathbf{I}^2$. *We define*

$$M_d(u_1, ..., u_d) = \min(u_1, ..., u_d), \tag{1.3}$$

and

$$W_d(u_1, ..., u_d) = \max(\sum_{i=1}^{d} u_i - (d-1), 0). \tag{1.4}$$

Let C' *be any d-subcopula with domain* S. *Then*

$$W_d(u_1, ..., u_d) \leq C'(u_1, ..., u_d) \leq M_d(u_1, ..., u_d), \tag{1.5}$$

for every $(u_1, ..., u_d) \in S$.
M_d *and* W_d *are known as the* ***Frechet-Hoeffding's bounds***.

M_d is always a d-copula for every $d \geq 2$ and W_d is a copula for $d = 2$, but W_d is not a copula for $d \geq 3$. The left side of (1.5) is best possible, in the sense that for every $d \geq 3$ and for every $(u_1, ..., u_d) \in \mathbf{I}^d$, there exists a d-copula C such that $C(u_1, ..., u_d) = W_d(u_1, ..., u_d)$.

Definition 1.1.5. *Let $(u_1, ..., u_d) \in \mathbf{I}^d$. We define the* ***product copula****, denoted as Π_d, by*

$$\Pi_d(u_1, ..., u_d) = \prod_{i=1}^{d} u_i. \tag{1.6}$$

The following theorem characterizes the independence between continuous random variables via copulas.

Theorem 1.1.6. *Let $X_1, ..., X_d$ be continuous random variables with unique d-copula C. Then $X_1, ..., X_d$ are independent if only if $C = \Pi_d$.*

If C is a d-copula, then by theorem 1.1.2 it is uniformly continuous, and hence by the Lebesgue's decomposition theorem we have that for every $(u_1, ..., u_d) \in \mathbf{I}^d$

$$C(u_1, ..., u_d) = A_C(u_1, ..., u_d) + S_C(u_1, ..., u_d),$$

where A_C is the absolutely continuous component with respect to the Lebesgue measure in $\mathbb{R}^d$ and

$$A_C(u_1, ..., u_d) = \int_0^{u_1} \dots \int_0^{u_d} \frac{\partial^d}{\partial x_1 \dots \partial x_d} C(x_1, ..., x_d) dx_d \dots dx_1;$$

and S_C is the singular component with respect to the Lebesgue measure in $\mathbb{R}^d$ and

$$S_C(u_1, ..., u_d) = C(u_1, ..., u_d) - A_C(u_1, ..., u_d).$$

If $C = A_C$ on $\mathbf{I}^d$, we say that C is absolutely continuous and its density is given by

$$c(u_1, ..., u_d) = \frac{\partial^d C(u_1, ..., u_d)}{\partial u_1 \cdots \partial u_d};$$

if $C = S_C$ on $\mathbf{I}^d$, we say that C is singular and we have that

$$\frac{\partial^d C(u_1, ..., u_d)}{\partial u_1 \cdots \partial u_d} = 0 \text{ a.s. } [\mathbb{P}_C];$$

and in any other case we will say that C has an absolutely continuous component A_C and a singular component S_C. It is common that, in this case, C is called hybrid.

If H is the distribution function related to C, as in Sklar's Theorem, we know that if the support of H is denoted by S_H, then

$$S_H = \left(\bigcup A\right)^c \text{ such that } A \text{ is an open set in } \mathbb{R}^d \text{ and } \mathbb{P}_H(A) = 0,$$

where $\mathbb{P}_H$ denotes the probability measure induced by H on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}))$.

We remember that the arbitrary union of open sets is an open set and its complement is a closed measurable set. Then, if S_C denotes the support of C we have that S_C is a closed set and

$$S_C = \left(\bigcup B\right)^c \text{ such that } B \text{ is an open set in } \mathbf{I}^d \text{ and } \mathbb{P}_C(B) = 0.$$

We can determine if a d-copula is singular via the Lebesgue's measure in $\mathbb{R}^d$, according with the following result: C is singular if only if the support of C has Lebesgue's measure zero.

We observe that the product copula Π_d is absolutely continuous, because for every $(u_1, ..., u_d) \in \mathbf{I}^d$

$$\begin{aligned} A_{\Pi_d}(u_1, ..., u_d) &= \int_0^{u_1} \cdots \int_0^{u_d} \frac{\partial^d}{\partial u_1 \cdots \partial u_d} \Pi_d(v_1, ..., v_d) dv_d \cdots dv_1 \\ &= \int_0^{u_1} \cdots \int_0^{u_d} 1 dv_d \cdots dv_1 \\ &= \prod_{i=1}^{d} u_i \\ &= \Pi_d(u_1, ..., u_d). \end{aligned} \tag{1.7}$$

Let M_d as in (1.3). The support of M_d, denoted by S_{M_d} is the main diagonal, i.e., $S_{M_d} = \{u_1 = u_2 = \cdots = u_{d-1} = u_d | u_i \in \mathbf{I} \text{ for every } i \in I_d\}$. Then M_d is singular, because $\lambda^d(S_{M_d}) = 0$, where λ^d is the Lebesgue's measure on $\mathbb{R}^d$. Besides $\partial^d M_d / \partial u_1 \cdots \partial u_d = 0$ a.s. $[\lambda^d]$.

Let $0 \leq \alpha \leq 1$ and $0 \leq \beta \leq 1$. We define for every $(u, v) \in \mathbf{I}^2$ the family of copulas

$$C_{\alpha,\beta}(u, v) = \min(u^{1-\alpha} v, u v^{1-\beta}) = \begin{cases} u^{1-\alpha} v & \text{if } u^\alpha \geq v^\beta; \\ u v^{1-\beta} & \text{if } u^\alpha \leq v^\beta. \end{cases}$$

This family is known as Marshall-Olkin family or Generalized Cuadras-Augé family. This family of copulas is hybrid, i.e., it has absolutely continuous and singular components. This follows because

$$\frac{\partial^2}{\partial u \partial v} C_{\alpha,\beta}(u, v) = \begin{cases} (1-\alpha) u^{-\alpha} & \text{if } u^\alpha > v^\beta; \\ (1-\beta) v^{-\beta} & \text{if } u^\alpha < v^\beta. \end{cases}$$

Then,if $A_{\alpha,\beta}$ denotes the absolutely continuous part of $C_{\alpha,\beta}$, we have that for $u^\alpha < v^\beta$

$$A_{\alpha,\beta}(u,v) = uv^{1-\beta} - \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}(u^\alpha)^{(\alpha+\beta-\alpha\beta)/\alpha\beta},$$

and for $u^\alpha > v^\beta$

$$A_{\alpha,\beta}(u,v) = u^{1-\alpha}v - \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}(v^\beta)^{(\alpha+\beta-\alpha\beta)/\alpha\beta},$$

i.e.,

$$A_{\alpha,\beta}(u,v) = C_{\alpha,\beta}(u,v) - \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}(\min(u^\alpha, v^\beta))^{(\alpha+\beta-\alpha\beta)/\alpha\beta}.$$

And if $S_{\alpha,\beta}$ denotes the singular part of $C_{\alpha,\beta}$ we have that for $u^\alpha = v^\beta$

$$S_{\alpha,\beta}(u,v) = \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}(\min(u^\alpha, v^\beta))^{(\alpha+\beta-\alpha\beta)/\alpha\beta}.$$

1.2 Archimedean copulas

In this section we introduce an important class of copulas called Archimedean. The Archimedean class has many properties and the copulas which belong to this class are easy to construct. Besides, many of the most used copulas belong to this family.

Definition 1.2.1. *Let φ be a continuous, strictly decreasing function such that $\varphi : \mathbf{I} \to [0,\infty]$ and $\varphi(1) = 0$. We define the* ***pseudo-inverse*** *of φ, denoted by $\varphi^{[-1]}$, as the function with domain $[0,\infty]$ and range $\mathbf{I}$ given by*

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t) & \text{if} \quad 0 \le t \le \varphi(0); \\ 0 & \text{if} \quad \varphi(0) \le t \le \infty, \end{cases}$$

where φ^{-1} is the usual inverse of φ.

Observe that if $\varphi(0) = \infty$, then $\varphi^{[-1]} = \varphi^{-1}$.

Theorem 1.2.2. *Let φ and $\varphi^{[-1]}$ as in Definition 1.2.1. Let $C : \mathbf{I}^2 \to \mathbf{I}$ be a function such that*

$$C(u,v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)). \tag{1.8}$$

Then C is a 2-copula if and only if φ is convex.

Definition 1.2.3. *If a 2-copula C can be represented as in* (1.8)*, then it is called an* ***Archimedean copula****; and the funtion φ is known as the* ***Archimedean generator*** *of C. If $\varphi^{[-1]} = \varphi^{-1}$, then φ is called a strict generator. In any other case φ is known as a non strict generator.*

We will see two examples of Archimedean copulas:

Example 1.2.4. *i) Let $\varphi(t) = -ln(t)$, for $t \in \mathbf{I}$. We notice that $\varphi(0) = \infty$, then $\varphi^{[-1]} = \varphi^{-1} =$* $\exp(-t)$ *and generates a copula. According to* (1.8) *the copula is given by*

$$C(u, v) = \exp(-[-ln(u) - ln(v)]) = uv = \Pi_2(u, v).$$

ii) Let $\varphi(t) = \frac{1}{\theta}\left(t^{-\theta} - 1\right)$, with $\theta \in [-1, \infty) \setminus \{0\}$. We observe that $\varphi(0) = \infty$, and then it follows that $\varphi^{[-1]} = \varphi^{-1} = (1 + \theta t)^{-1/\theta}$. Then φ generates a copula given by

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

C is known as the Clayton copula.

Definition 1.2.5. *A function g is said to be* ***completely monotonic*** *on an interval J if and only if it is continuous and satisfies*

$$(-1)^k \frac{d^k}{dt^k} g(t) \geq 0, \tag{1.9}$$

for all t in the interior of J and for every $k = 0, 1, 2, ...$

Theorem 1.2.6. *Let $\varphi : \mathbf{I} \to [0, \infty]$ be a continuous, strictly decreasing function such that $\varphi(1) = 0$ and $\varphi(0) = \infty$. Let $d \geq 2$ and $C : \mathbf{I}^d \to \mathbf{I}$ be a function given by*

$$C(u_1, ..., u_d) = \varphi^{-1}(\varphi(u_1) + \cdots + \varphi(u_d)). \tag{1.10}$$

Then C is a d-copula for every $d \geq 2$ if and only if φ^{-1} is completely monotonic on $[0, \infty)$.

A copula C that satisfies (1.10) is called **Archimedean d-copula**, for every $d \geq 2$.

Example 1.2.7. *Let $\theta > 0$. We define $\varphi_\theta(t) = \frac{1}{\theta}(1/t^\theta - 1)$. Then, clearly, φ_θ satisfies the condition of theorem 1.2.6. We note that $\varphi_\theta^{-1} = (1 + \theta t)^{-1/\theta}$ and*

$$(-1)^k \frac{d^k}{dt^k} \varphi_\theta^{-1} = (-1)^{2k} \frac{(1 + \theta t)^{-(1+k\theta)/\theta}}{\theta^k} \prod_{i=1}^{k-1} (1 + (i-1)\theta).$$

Hence

$$C_\theta(u_1, ..., u_d) = \left(u_1^{-\theta} + \cdots + u_d^{-\theta} - n + 1\right)^{-1/\theta},$$

is an Archimedean d-copula.
This is known as the Clayton family of d-copulas for $\theta > 0$.

1.3 Dependence

In this section we present the definition of one measure of dependence: the concordance. We also give the definition of two measures of association: Kendall's tau and Spearman's rho, and we discuss their relation with copulas.

Definition 1.3.1. *Let (x_i, y_i) and (x_j, y_j) denote two observations from a vector (X, Y) of continuous random variables. We say that (x_i, y_i) and (x_j, y_j) are* ***concordant*** *if and only if $x_i < x_j$ and $y_i < y_j$, or if $x_i > x_j$ and $y_i > y_j$. In the same way, we say that (x_i, y_i) and (x_j, y_j) are* ***disconcordant*** *if and only if $x_i < x_j$ and $y_i > y_j$, or if $x_i < x_j$ and $y_i > y_j$.*

Definition 1.3.2. *Let $\{(x_1, y_1), ..., (x_n, y_n)\}$ be a random sample of size n from a vector (X, Y) of continuous random variables. Let c be the number of concordant pairs and d be the number of discordant pairs. We define the* ***sample version of the Kendall's tau****, denoted by τ, as*

$$\tau = \frac{c-d}{c+d} = \frac{c-d}{\begin{pmatrix} n \\ 2 \end{pmatrix}}$$

i.e., τ is the probability of concordance minus the probability of discordance.

Following the same idea, the population version of Kendall's tau is given by the difference between the probability of concordance and the probability of discordance.

Definition 1.3.3. *Let (X_1, Y_1) and (X_2, Y_2) be independent and identically distributed random vectors, each with joint distribution function H and copula C. We define the* ***population version of Kendall's tau****, denoted by $\tau_{X,Y}$ or by τ_C, as*

$$\tau_{X,Y} = \tau_C = \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0]. \tag{1.11}$$

The next theorem gives a way to calculate the population Kendall's tau from continuous random vectors.

Theorem 1.3.4. *Let* (X_1, Y_1) *and* (X_2, Y_2) *be independent vectors of continuous random variables with joint distributions* H_1 *and* H_2*, respectively, with common margins* F *(of* X_1*and* X_2*) and* G *(of* Y_1 *and* Y_2*). Let* C_1 *and* C_2 *be the copulas of* (X_1, Y_1) *and* (X_2, Y_2)*, respectively. Let* Q *denote the difference between the probabilities of concordance and discordance of* (X_1, Y_1) *and* (X_2, Y_2)*, i.e.,*

$$Q = Q(C_1, C_2) = \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0]. \tag{1.12}$$

Then

$$Q(C_1, C_2) = 4 \int\int_{\mathbf{I}^2} C_2(u, v) dC_1(u, v) - 1.$$

As a direct consequence of the theorem 1.3.4, we have that if $\tau_{X,Y}$ is defined as in (1.11), then

$$\tau_{X,Y} = \tau_C = Q(C, C) = 4 \int\int_{\mathbf{I}^2} C(u, v) dC(u, v) - 1 = 4\mathbb{E}[C(u, v)] - 1.$$

Now, we give the definition of another measure of association based on concordance and discordance, called Spearman's rho.

Definition 1.3.5. *Let* (X_1, Y_1)*,* (X_2, Y_2) *and* (X_3, Y_3) *be three independent random vectors with a common joint distribution function* H *and copula* C*. We define the* ***population version of Spearman's rho****, denoted by* $\rho_{X,Y}$ *or by* ρ_C*, as*

$$\rho_{X,Y} = \rho_C = 3(\mathbb{P}[(X_1 - X_2)(Y_1 - Y_3) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_3) < 0]), \tag{1.13}$$

The Spearman's rho is defined as the probability of concordance minus the probability of discordance for a pair of vectors with the same margins, but one has distribution function H, while the other vector has independent components.

In the same way that Theorem 1.3.4, the next theorem establishes a way to calculate the Spearman's rho for random continuous variables.

Theorem 1.3.6. *Let* X *and* Y *be continuous random variables with copula* C*. Then the population version of the Spearman's rho, defined as in* (1.13)*, is given by*

$$\rho_{X,Y} = \rho_C = 3Q(C, \Pi_2) = 12 \int\int_{\mathbf{I}^2} uv dC(u, v) - 3 = 12 \int\int_{\mathbf{I}^2} C(u, v) du dv - 3. \tag{1.14}$$

Let X and Y be continuous random variables with distribution functions F and G, respectively, and copula C. If $U = F(X)$ and $V = G(Y)$, then U and V have the same distribution, uniform

$(0,1)$, and joint distribution C. Then $\mathbb{E}[U] = \mathbb{E}[V] = 1/2$ and $Var[U] = Var[V] = 1/12$, and this implies

$$\rho_{X,Y} = \rho_C = 12 \int\int_{\mathbf{I}^2} uv dC(u,v) - 3 = 12\mathbb{E}[UV] - 3 = \frac{\mathbb{E}[UV] - \frac{1}{4}}{\frac{1}{12}} = \frac{\mathbb{E}[UV] - \mathbb{E}[U]\mathbb{E}[V]}{\sqrt{Var[U]}\sqrt{Var[V]}}, \tag{1.15}$$

i.e., the population version of Spearman's rho for X and Y is equal to Pearson's correlation coefficient for U and V.

Chapter 2

Sample-d copula of order m

In this chapter we present the definition and the most important results about the sample d-copula of order m. Moreover, in the first section, we establish two almost unknown results about empirical distribution functions and empirical copulas. In the third section we compare the sample d-copula of order m with the empirical copula, and we will see that the sample copula is a better approximation in many senses. Finally, in the fourth section, we give the definition of the copula called the checkerboard approximation and we give some basic result; besides we establish a Glivenko-Cantelli Theorem for the sample copula.

2.1 Empirical functions

In this section we present the theorems of Glivenko-Cantelli for the empirical distribution function and for sample copulas. Besides, we give two results that establish bounds in the opposite way that the Glivenko-Cantelli theorem, we can think these results as “anti-Glivenko-Cantelli Theorems”. It is surprising that these results have been little studied previously.

Definition 2.1.1. *Let $X_1, ..., X_n$ be a random sample of size n from a continuous random variables X and let $X_{(1)}, ..., X_{(n)}$ be their order statistics. The **rank function** $r : I_n \times \mathbb{R}^n \to I_n$ is defined by*

$$r(j, X_1, ..., X_n) = k, \text{ if and only if } X_j = X_{(k)} \text{ where } j, k \in I_n.$$

Definition 2.1.2. *Let $\underline{X}_1, ..., \underline{X}_n$ be a random sample of size n from a continuous random vector $\underline{X}$ of dimension d, where $\underline{X}_i = (X_{i,1}, ..., X_{i,d}) \in \mathbb{R}^d$, for every $i \in I_n$. Let $i \in I_n$, the i-th **modified sample** $\underline{Y}_i = (Y_{i,1}, ..., Y_{i,d})$, is defined by*

$$Y_{i,j} = \frac{1}{n} r(i, X_{1,j}, ..., X_{n,j}) \text{ for every } j \in I_d.$$

Example 2.1.3. *Let $\underline{X}_1 = (2.2178, 2.6011)$, $\underline{X}_2 = (-2.1351, 1.9449)$, $\underline{X}_3 = (0.1139, 0.2113)$, $\underline{X}_4 = (-0.3874, 3.0680)$ and $\underline{X}_5 = (0.7394, -2.0514)$ be a random sample of size $n = 5$ from a* *bivariate normal distribution with mean* $\mu = (0,0)$ *and correlation coefficient* $\rho = 0.4$. *Then*

$$\underline{Y}_1 = \left(\frac{5}{5}, \frac{4}{5}\right), \underline{Y}_2 = \left(\frac{1}{5}, \frac{3}{5}\right), \underline{Y}_3 = \left(\frac{3}{5}, \frac{2}{5}\right), \underline{Y}_4 = \left(\frac{2}{5}, \frac{5}{5}\right) \text{ and } \underline{Y}_5 = \left(\frac{4}{5}, \frac{1}{5}\right).$$

Definition 2.1.4. *Let* $\underline{X}_1, ..., \underline{X}_n$ *be a random sample of size* n *from a random vector* $\underline{X}$ *of dimension* d*, with continuous joint distribution* H *and unique copula* C*. Let* $\underline{Y}_1, ..., \underline{Y}_n$ *be the corresponding modified sample. We define the* ***empirical copula****, denoted by* $C_n : \mathbf{I}^d \to \mathbf{I}$*, by*

$$C_n(u_1, ..., u_d) = \frac{1}{n}\sum_{i=1}^{n} 1_{(-\infty,u_1]\times\cdots\times(-\infty,u_d]}(Y_{i,1}, ..., Y_{i,d}), \text{ for every } (u_1, ..., u_d) \in \mathbf{I}^d. \tag{2.1}$$

The following Theorem is the version of the Glivenko-Cantelli's Theorem for empirical copulas:

Theorem 2.1.5. *(Glivenko-Cantelli) Let* C_n *be the empirical copula constructed from a sample of size* n *from a continuous joint distribution* H *with copula* C*. Then*

$$\lim_{n\to\infty} \sup_{(u_1,...,u_d)\in\mathbf{I}^d} |C_n(u_1, ..., u_d) - C(u_1, ..., u_d)| = 0 \text{ a.s. } [\mathbb{P}_C],$$

where $[\mathbb{P}_C]$ *is the probability measure on* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ *induced by the* d*-copula* C*.*

Definition 2.1.6. *Let* $X_1, ..., X_n$ *be a random sample of size* n *from a random variable* X *with distribution function* F*. We define, for every* $x \in \mathbb{R}$*, the* ***empirical distribution function****, denoted by* F_n*, as*

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{(-\infty,x]}(X_i). \tag{2.2}$$

A well known result for distribution functions is the Glivenko-Cantelli's Theorem:

Theorem 2.1.7. *(Glivenko-Cantelli) Let* $X_1, ..., X_n$ *be a random sample of size* n *from a random variable* X *with distribution function* F*, and let* F_n *be the empirical distribution function as in* (2.2)*. Then*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |F_n(x) - F(x)| = 0 \text{ a.s. } [\mathbb{P}_F].$$

where $\mathbb{P}_F$ *is the probability measure induced by* F *on* $\mathbb{R}$*.*

The Glivenko-Cantelli's Theorem indicates that when the sample size goes to infinity the empirical distribution converges a.s. to the theoretical distribution. But, what happens for a fixed n? What is the bound for the "worst" sample? An answer to this question is given by the almost unknown result:

Lemma 2.1.8. *Let $X_1, ..., X_n$ be a random sample of size n from a random variable X with continuous distribution function F. Let F_n be the empirical distribution function as in* (2.2). *Then*

$$\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \geq \frac{1}{2n} \text{ a.s. } [\mathbb{P}_F]. \tag{2.3}$$

Proof: Assume that F is continuous and that the sample is ordered, that is, $X_1 < X_2 < \cdots < X_n$. As F_n is a step function, we have for every $k \in \{1, ..., n\}$

$$F_n(X_k) = \frac{k}{n} \text{ and } F_n(X_k^-) = \frac{k-1}{n}.$$

Then

$$\begin{aligned}\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| &= \max_{1\leq k\leq n} (\max(|F_n(X_k^-) - F(X_k)|, |F_n(X_k) - F(X_k)|)) \\ &= \max_{1\leq k\leq n} \left(\max \left(\left|\frac{k-1}{n} - F\left(X_k\right)\right|, \left|\frac{k}{n} - F\left(X_k\right)\right| \right) \right).\end{aligned} \tag{2.4}$$

If we also assume that, for every $k \in \{1, ..., n\}$, $F(X_k) = (2k-1)/(2n)$, that is, $X_k = F^{(-1)}((2k-1)/(2n))$, then

$$\left|\frac{k-1}{n} - \frac{2k-1}{2n}\right| = \left|\frac{2k-2-2k+1}{2n}\right| = \frac{1}{2n} = \left|\frac{2k-(2k-1)}{2n}\right| = \left|\frac{k}{n} - \frac{2k-1}{2n}\right|.$$

If $F(X_k) \neq (2k-1)/(2n)$ then

$$\max_{1\leq k\leq n} \left(\left|\frac{k-1}{n} - F(X_k)\right|, \left|\frac{k}{n} - F(X_k)\right| \right) > \frac{1}{2n}.$$

Hence

$$\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \geq \max_{1\leq k\leq n} \frac{1}{2n} = \frac{1}{2n},$$

and (2.3) is satisfied.

In the same sense that in the last lemma, we establish a bound between the real copula and the empirical copula for the supremum distance and for a fixed sample size.

Lemma 2.1.9. *Let $\underline{X}_1, ..., \underline{X}_n$ be a random sample of size n from a random vector $\underline{X}$ of dimension d, with continuous joint distribution function H and copula C. Let $\underline{Y}_1, ..., \underline{Y}_n$ be the corresponding modified sample, and let C_n be the empirical copula. Then*

$$\sup_{(u_1,...,u_d)\in\mathbf{I}^d} |C_n(u_1, ..., u_d) - C(u_1, ..., u_d)| \geq \frac{1}{n} \text{ a.s. } [\mathbb{P}_C]. \tag{2.5}$$

Proof: Let $0 < \varepsilon < 1/n$. Since C is a d-copula, then $C(\varepsilon, 1, ..., 1) = \varepsilon$, but by definition of the empirical copula we have that $C_n(\varepsilon, 1, ..., 1) = 0$, because we can not see any points with any coordinates less than or equal to ε. Then

$$\lim_{\varepsilon \uparrow (1/n)} |C_n(\varepsilon, 1, ..., 1) - C(\varepsilon, 1, ..., 1)| = \frac{1}{n},$$

and the result follows.

2.2 Sample d-copula of order m

The first concept presented in this section is the generalized transformation matrix. The matrices of this kind are generated by a probability measure.

The main topic of the section is the sample d-copula of order m; this copula was first introduced in [9], and the article [11] presents some improvements. The sample copula assigns uniform mass to every d-box generated by a generalized transformation matrix that is determined by the sample.

We establish some important results: given a fixed size for the sample, the partition generated by the generalized transformation is independent from the sample; the density of the sample is constant on every d-box generated by the partition, and the sample copula is in fact a copula.

Definition 2.2.1. *Let $I_n = \{1, 2, ..., n\}$. For dimension $d \geq 2$, let $m \in \mathbb{N}$, we define $I_m^d = \times_{i=1}^d I_m$. Let τ be a probability measure on $(I_m^d, 2^{I_m^d})$, τ is known as a* ***generalized transformation matrix*** *if for all $j \in I_d$ and for all $k \in I_m$*

$$\sum_{\underline{i} \in I_m^d, i_j = k} \tau(\underline{i}) > 0,$$

where $\underline{i} = (i_1, ..., i_{j-1}, i_j = k, i_{j+1}, ..., i_d) \in I_m^d$. τ can be thought of as a d-dimensional matrix τ, considering

$$\tau(\underline{i}) = \tau_{i_1,...,i_d} \text{ if } \underline{i} = (i_1, ..., i_d) \in I_m^d.$$

Example 2.2.2. *Let*

$$A = \begin{pmatrix} 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix},$$

and

$$B = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}.$$

Then A is a generalized transformation matrix and B is not.

Definition 2.2.3. *Let $\tau = (\tau_{i,j})_{i,j\in\{1,\dots,m\}}$ be a generalized transformation matrix where $d = 2$. Define $\{q_{1,0}, q_{1,1}, ..., q_{1,m}\}$ and $\{q_{2,0}, q_{2,1}, ..., q_{2,m}\}$ two partitions of $\mathbf{I}$, such that $q_{1,0} = q_{2,0} = 0$ and for $i, j \in I_m$ we have that*

$$q_{1,i} = \sum_{i'=1}^{i} \sum_{j\in I_m} \tau_{i',j} \text{ and } q_{2,j} = \sum_{j'=1}^{j} \sum_{i\in I_m} \tau_{i,j'}.$$

We also define the partition induced by τ on $\mathbf{I}^2$ by

$$Q^m_{i,j} = \langle q_{1,i-1}, q_{1,i}] \times \langle q_{2,j-1}, q_{2,j}] \text{ for every } (i, j) \in I_m \times I_m,$$

where the $\langle$ notation indicates that the left end of the interval is closed if $i = 1$ or $j = 1$, and open in any other case. Let Π_2 be the product 2-copula, and define the $\tau(\Pi_2)$ transformation by

$$\begin{aligned}\tau(\Pi_2)(u, v) = &\sum_{i'<i,j'<j} \tau_{i',j'} + \frac{u - q_{1,i-1}}{q_{1,i} - q_{1,i-1}} \sum_{j'<j} \tau_{i,j'} + \frac{v - q_{2,j-1}}{q_{2,j} - q_{2,j-1}} \sum_{i'<i} \tau_{i',j} \\ &+ \tau_{i,j}\Pi_2\left(\frac{u - q_{1,i-1}}{q_{1,i} - q_{1,i-1}}, \frac{v - q_{2,j-1}}{q_{2,j} - q_{2,j-1}}\right),\end{aligned} \tag{2.6}$$

with $u, v \in Q^m_{i,j}$ for every $i, j \in I_m$.

Equation (2.6) is the same as equation 2.3.2 in Lemma 2.3.5 in the proof of Sklar's Theorem in [18], using the subcopula generated by the generalized transformation matrix τ. Equation (2.6) is a bilinear interpolation and hence $\tau(\Pi_2)$ assigns the mass uniformly in every 2-box $Q_{i,j}$.

Definition 2.2.4. *Let $m \geq 2$ and let $\tau = (\tau_{i_1,\dots,i_d})_{(i_1,\dots,i_d)\in(I_m)^d}$ be a generalized transformation matrix. We define $q_{1,0} = q_{2,0} = \cdots = q_{d,0} = 0$, and for every $j \in I_d$ and for every $k \in I_m$*

$$q_{j,k} = \sum_{i_j=1}^{k} \sum_{i_1=1}^{m} \cdots \sum_{i_{j-1}=1}^{m} \sum_{i_{j+1}=1}^{m} \cdots \sum_{i_d=1}^{m} \tau_{i_1,\dots,i_{j-1},i_j,i_{j+1},\dots,i_d}.$$

Then $0 = q_{j,0} < q_{j,1} < \cdots < q_{j,m-1} < q_{j,m} = 1$ is a partition of $\mathbf{I}$, induced by the matrix τ in the j-coordinate. For every $\underline{i} = (i_1, ..., i_d) \in (I_m)^d$ we define

$$Q^m_{\underline{i}} = \langle q_{1,(i_1-1)}, q_{1,i_1}] \times \langle q_{2,(i_2-1)}, q_{2,i_2}] \times \cdots \times \langle q_{d,(i_d-1)}, q_{d,i_d}]. \tag{2.7}$$

Then the family $(Q^m_{\underline{i}})_{\underline{i}\in(I_m)^d}$ is a partition of $\mathbf{I}^d$.

Definition 2.2.5. *Let* $2 \leq m \leq n$ *and let* $\underline{X}_1, ..., \underline{X}_n$ *be a random sample of a size* n *from a random vector* $\underline{X}$ *of dimension* d*, with continuous joint distribution* H *or* d*-copula* C*, where* $\underline{X}_i = (X_{i,1}, ..., X_{i,d}) \in \mathbb{R}^d$*, for every* $i = 1, ..., n$*. Let* $U_n = \{\underline{Y}_1, ..., \underline{Y}_n\}$ *be the corresponding modified sample. Define the uniform partition of size* m *of* $\mathbf{I}^d$*, where for every* $\underline{i} = (i_1, ..., i_d) \in (I_m)^d$

$$R_{\underline{i}}^m = \left\langle \frac{i_1 - 1}{m}, \frac{i_1}{m} \right] \times \cdots \times \left\langle \frac{i_d - 1}{m}, \frac{i_d}{m} \right]. \tag{2.8}$$

Define

$$s_{i_1,...,i_d}^{n,(m)} = \frac{card(R_{\underline{i}}^m \cap U_n)}{n}, \tag{2.9}$$

where $card(\cdot)$ *denotes de cardinality of a set. Let*

$$S_m^n = (s_{i_1,...,i_d}^{n,(m)})_{(i_1,...,i_d) \in (I_m)^d}, \tag{2.10}$$

then S_m^n *is always a* d*-dimensional generalized transformation matrix. Finally, let* $(Q_{\underline{i}}^m)_{\underline{i} \in (I_m)^d}$ *be the partition of* $\mathbf{I}^d$ *induced by the generalized transformation matrix* S_m^n *given in equation* (2.7)*. Using the partition* $(Q_{\underline{i}}^m)_{\underline{i} \in (I_m)^d}$*, we define the* ***sample*** d***-copula of order*** m *by*

$$C_m^n(u_1, ..., u_d) = S_m^n(\Pi_d)(u_1, ..., u_d), \tag{2.11}$$

as in the generalization of equation (2.6)*, where* Π_d *is the product copula in* $\mathbf{I}^d$*.*

The function C_m^n was first proposed in [9]. We will see that C_m^n is an estimator of the true copula C because is an estimator of $C^{(m)}$, the Checkerboard approximation, see Definition 2.4.2 below. For a more in-depth study of C_m^n see [11].
We give the following example to clarify the definition of the Sample Copula of Order m:

Example 2.2.6. *Let* $\underline{X}_1 = (-0.2787191, 0.8874746)$*,* $\underline{X}_2 = (-1.60796965, 0.9300367)$*,* $\underline{X}_3 = (3.85470838, -2.7634594)$*,* $\underline{X}_4 = (3.83099590, -1.7714260)$ *and* $\underline{X}_5 = (-0.87848834, -0.78799474)$ *be a random sample of size* $n = 5$ *from a bivariate* t *distribution with 3 degrees of freedom, mean* $\mu = (0, 0)$ *and variance-covariance matrix*

$$V = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}.$$

According to definition 2.1.2, the modified sample is given by $\underline{Y}_1 = (3/5, 4/5), \underline{Y}_2 = (1/5, 5/5), \underline{Y}_3 = (5/5, 1/5), \underline{Y}_4 = (4/5, 2/5), \underline{Y}_5 = (2/5, 3/5)$*. Let* $U_5 = \{\underline{Y}_1, ..., \underline{Y}_5\}$*.*

For $m = 2$ *we have*
$R^2_{1,1} \cap U_5 = \emptyset$,
$R^2_{1,2} \cap U_5 = \{\underline{Y}_2, \underline{Y}_5\}$,
$R^2_{2,1} \cap U_5 = \{\underline{Y}_3, \underline{Y}_4\}$,
$R^2_{2,2} \cap U_5 = \{\underline{Y}_1\}$.
Then the matrix defined in (2.10) *is given by*

$$S^5_2 = \begin{pmatrix} 0 & \frac{2}{5} \\ \frac{2}{5} & \frac{1}{5} \end{pmatrix}.$$

The partition induced by S^5_2*, as in definition 2.2.4, is determined by*
$q_{1,0} = q_{2,0} = 0, q_{1,1} = q_{2,1} = 2/5, q_{1,2} = q_{2,2} = 1$.
Then
$Q^2_{1,1} = [0, 2/5] \times [0, 2/5], Q^2_{1,2} = [0, 2/5] \times (2/5, 1], Q^2_{2,1} = (2/5, 1] \times [0, 2/5], Q^2_{2,2} = (2/5, 1] \times (2/5, 1]$.
By equation (2.11)*, the sample 2-copula of order 2 is given by*

$$C^5_2(u,v) = \begin{cases} 0 & \textit{if} \quad (u,v) \in Q^2_{1,1} \\ u\left(\frac{5v-2}{3}\right) & \textit{if} \quad (u,v) \in Q^2_{1,2} \\ \left(\frac{5u-2}{3}\right)v & \textit{if} \quad (u,v) \in Q^2_{2,1} \\ \frac{2}{5}\left(\frac{5(u+v)-4}{3}\right) + \frac{1}{5}\left(\frac{5u-2}{3}\right)\left(\frac{5v-2}{3}\right) & \textit{if} \quad (u,v) \in Q^2_{2,2} \end{cases},$$

and its density is given by

$$c^5_2(u,v) = \begin{cases} 0 & \textit{if} \quad (u,v) \in Q^2_{1,1} \\ \frac{5}{3} & \textit{if} \quad (u,v) \in Q^2_{1,2} \cup Q^2_{2,1} \\ \frac{5}{9} & \textit{if} \quad (u,v) \in Q^2_{2,2}. \end{cases}$$

We observe that the density is constant in every 2-box and is given by $s_{i,j}/\lambda^2(Q^2_{i,j})$*, for* $i, j \in I_m$*, where* λ^2 *denotes the Lebesgues measure in* $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$.
Now, for $m = 3$ *we have*
$R^3_{i,j} \cap U_5 = \emptyset$*, for* $(i,j) \in \{(1,1), (1,2), (2,1), (3,3)\}$,
$R^3_{1,3} \cap U_5 = \{\underline{Y}_2\}$,
$R^3_{2,2} \cap U_5 = \{\underline{Y}_5\}$,
$R^3_{2,3} \cap U_5 = \{\underline{Y}_1\}$,
$R^3_{3,1} \cap U_5 = \{\underline{Y}_3\}$,

$R^3_{3,2} \cap U_5 = \{\underline{Y}_4\}$.

Then

$$S^5_3 = \begin{pmatrix} 0 & 0 & \frac{1}{5} \\ 0 & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & 0 \end{pmatrix}.$$

The partition induced by S^5_3 *is determined by*

$q_{1,0} = q_{2,0} = 0, q_{1,1} = q_{2,1} = 1/5, q_{1,2} = q_{2,2} = 3/5, q_{1,3} = q_{2,3} = 1.$

Then

$Q^3_{1,1} = [0,1/5] \times [0,1/5], Q^3_{1,2} = [0,1/5] \times (1/5,3/5], Q^3_{1,3} = [0,1/5] \times (2/5,1], Q^3_{2,1} = (1/5,3/5] \times [0,1/5], Q^3_{2,2} = (1/5,3/5] \times (1/5,3/5], Q^3_{2,3} = (1/5,3/5] \times (3/5,1], Q^3_{3,1} = (3/5,1] \times [0,1/5], Q^3_{3,2} = (3/5,1] \times (1/5,3/5], Q^3_{3,3} = (3/5,1] \times (3/5,1].$

Hence, the sample 2-copula of order 3 is given by

$$C^5_3(u,v) = \begin{cases} 0 & \text{if} \quad (u,v) \in Q^3_{1,1} \cup Q^3_{1,2} \cup Q^3_{2,1} \\ u\left(\frac{5v-3}{2}\right) & \text{if} \quad (u,v) \in Q^3_{1,3} \\ \frac{1}{5}\left(\frac{5u-1}{2}\right)\left(\frac{5v-1}{2}\right) & \text{if} \quad (u,v) \in Q^3_{2,2} \\ \frac{1}{5}\left(\frac{5(u+v)-4}{2}\right) + \frac{1}{5}\left(\frac{5u-1}{2}\right)\left(\frac{5v-3}{2}\right) & \text{if} \quad (u,v) \in Q^3_{2,3} \\ \left(\frac{5u-3}{2}\right)v & \text{if} \quad (u,v) \in Q^3_{3,1} \\ \frac{1}{5}\left(\frac{5(u+v)-4}{2}\right) + \frac{1}{5}\left(\frac{5u-3}{2}\right)\left(\frac{5v-1}{2}\right) & \text{if} \quad (u,v) \in Q^3_{3,2} \\ u+v-1 & \text{if} \quad (u,v) \in Q^3_{3,3}. \end{cases}$$

and its density is given by

$$c^5_3(u,v) = \begin{cases} 0 & \text{if} \quad (u,v) \in Q^3_{1,1} \cup Q^3_{1,2} \cup Q^3_{2,1} \cup Q^3_{3,3} \\ \frac{5}{2} & \text{if} \quad (u,v) \in Q^3_{1,3} \cup Q^3_{3,1} \\ \frac{5}{4} & \text{if} \quad (u,v) \in Q^3_{2,2} \cup Q^3_{2,3}3 \cup Q^3_{3,2}. \end{cases}$$

Again, we observe that the density is constant in every 2-box and is given by $s_{i,j}/\lambda^2(Q^3_{i,j})$, *for* $i,j \in I_m$.

For $m = 4$ *we have*

$R^4_{i,j} \cap U_5 = \emptyset$, *for* $(i,j) \in \{(1,1),(1,2),(1,3),(2,1),(2,2),(2,4),(3,1),(3,2),(3,3),(4,3),(4,4)\}$,

$R^4_{1,4} \cap U_5 = \{\underline{Y}_2\}$,

$R^4_{3,4} \cap U_5 = \{\underline{Y}_1\}$,
$R^4_{4,1} \cap U_5 = \{\underline{Y}_3\}$,
$R^4_{4,2} \cap U_5 = \{\underline{Y}_4\}$.
Then

$$S^5_4 = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{5} \\ 0 & 0 & \frac{1}{5} & 0 \\ 0 & 0 & 0 & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & 0 & 0 \end{pmatrix}.$$

The partition induced by S^5_4 *is given by*
$q_{1,0} = q_{2,0} = 0, q_{1,1} = q_{2,1} = 1/5, q_{1,2} = q_{2,2} = 2/5, q_{1,3} = q_{2,3} = 3/5, q_{1,4} = q_{2,4} = 1.$
Then
$Q^4_{1,1} = [0,1/5] \times [0,1/5], Q^4_{1,2} = [0,1/5] \times (1/5,2/5], Q^4_{1,3} = [0,1/5] \times (2/5,3/5], Q^4_{1,4} = [0,1/5] \times (3/5,1], Q^4_{2,1} = (1/5,2/5] \times [0,1/5], Q^4_{2,2} = (1/5,2/5] \times (1/5,2/5], Q^4_{2,3} = (1/5,2/5] \times (2/5,3/5], Q^4_{2,4} = (1/5,2/5] \times (3/5,1], Q^4_{3,1} = (2/5,3/5] \times [0,1/5], Q^4_{3,2} = (2/5,3/5] \times (1/5,2/5], Q^4_{3,3} = (2/5,3/5] \times (2/5,3/5], Q^4_{3,4} = (2/5,3/5] \times (3/5,1], Q^4_{4,1} = (3/5,1] \times [0,1/5], Q^4_{4,2} = (3/5,1] \times (1/5,2/5], Q^4_{4,3} = (3/5,1] \times (2/5,3/5], Q^4_{4,4} = (3/5,1] \times (3/5,1].$
Thus, the sample 2-copula of order 4 is given by

$$C^5_4(u,v) = \begin{cases} 0 & \text{if} \quad (u,v) \in A \\ u\left(\frac{5v-3}{2}\right) & \text{if} \quad (u,v) \in Q^4_{1,4} \\ \left(\frac{5u-1}{5}\right)(5v-2) & \text{if} \quad (u,v) \in Q^4_{2,3} \\ \frac{1}{5}\left(5u-1+\frac{5v-3}{2}\right) & \text{if} \quad (u,v) \in Q^4_{2,4} \\ \frac{5u-2}{5} & \text{if} \quad (u,v) \in Q^4_{3,3} \\ \frac{1}{5}\left(1+\frac{5v-3}{2}+(5u-2)\left(\frac{5v-3}{2}\right)\right) & \text{if} \quad (u,v) \in Q^4_{3,4} \\ \left(\frac{5u-3}{2}\right)v & \text{if} \quad (u,v) \in Q^4_{4,1} \\ \frac{1}{5}\left(\frac{5u-3}{2}+\left(\frac{5u-3}{2}\right)(5v-1)\right) & \text{if} \quad (u,v) \in Q^4_{4,2} \\ u+v-1 & \text{if} \quad (u,v) \in Q^4_{4,3} \cup Q^4_{4,4}, \end{cases}$$

where $A = Q^4_{1,1} \cup Q^4_{1,2} \cup Q^4_{1,3} \cup Q^4_{2,1} \cup Q^4_{2,2} \cup Q^4_{3,1} \cup Q^4_{3,2}$.
The density is given by

$$c_4^5(u,v)=\begin{cases} 0 & if \quad (u,v)\in Q^4_{1,1}\cup Q^4_{1,2}\cup Q^4_{1,3}\cup Q^4_{2,1}\cup Q^4_{2,2}\cup Q^4_{2,4}\cup Q^4_{3,1}\cup Q^4_{3,2}\cup Q^4_{3,3}\cup Q^4_{4,3}\cup Q^4_{4,4}\\ \frac{5}{2} & if \quad (u,v)\in Q^4_{1,4}\cup Q^4_{3,4}\cup Q^4_{4,1}\cup Q^4_{4,2}\\ 5 & if \quad (u,v)\in Q^4_{2,3}. \end{cases}$$

We observe that the density is constant in every 2-box is constant and is given by $s_{i,j}/\lambda^2(Q^4_{i,j})$, for $i,j\in I_m$.

Let $m=5$. Then we have that

$R^5_{i,j}\cap U_5=\emptyset$, *for* $(i,j)\in I^2_m\setminus\{(1,5),(2,3),(3,4),(4,2),(5,1)\}$,

$R^5_{1,5}\cap U_5=\{\underline{Y}_2\}$,

$R^5_{2,3}\cap U_5=\{\underline{Y}_5\}$,

$R^5_{3,4}\cap U_5=\{\underline{Y}_1\}$,

$R^5_{4,2}\cap U_5=\{\underline{Y}_4\}$,

$R^5_{5,1}\cap U_5=\{\underline{Y}_3\}$.

Then the generalized transformation matrix is

$$S^5_5=\begin{pmatrix} 0 & 0 & 0 & 0 & \frac{1}{5}\\ 0 & 0 & \frac{1}{5} & 0 & 0\\ 0 & 0 & 0 & \frac{1}{5} & 0\\ 0 & \frac{1}{5} & 0 & 0 & 0\\ \frac{1}{5} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The matrix induces the partition

$q_{1,0}=q_{2,0}=0, q_{1,1}=q_{2,1}=1/5, q_{1,2}=q_{2,2}=2/5, q_{1,3}=q_{2,3}=3/5, q_{1,4}=q_{2,4}=4/5, q_{1,5}=q_{2,5}=1$,

and in this case we have that $(Q^5_{i,j})_{(i,j)\in I^2_5}=(R^5_{i,j})_{(i,j)\in I^2_5}$.

The sample 2-copula of order 5 is given by

$$C^5_5(u,v)=\begin{cases} 0 & if \quad (u,v)\in B\\ u(5v-4) & if \quad (u,v)\in Q^5_{1,5}\\ \frac{1}{5}(5u-1)(5v-2) & if \quad (u,v)\in Q^5_{2,3}\\ \frac{1}{5}(5u-1) & if \quad (u,v)\in Q^5_{2,4}\\ u+v-1 & if \quad (u,v)\in C\\ \frac{1}{5}(5v-2) & if \quad (u,v)\in Q^5_{3,3}\\ \frac{1}{5}(1+(5u-2)(5v-3)) & if \quad (u,v)\in Q^5_{3,4}\\ \frac{1}{5}(5u-3)(5v-1) & if \quad (u,v)\in Q^5_{4,2}\\ (5u-4)v & if \quad (u,v)\in Q^5_{5,1}, \end{cases}$$

where $B=Q^5_{1,1}\cup Q^5_{1,2}\cup Q^5_{1,3}\cup Q^5_{1,4}\cup Q^5_{2,1}\cup Q^5_{2,2}\cup Q^5_{3,1}\cup Q^5_{3,2}\cup Q^5_{4,1}$ *and* $C=Q^5_{2,5}\cup Q^5_{3,5}\cup Q^5_{4,3}\cup$

$Q^5_{4,4} \cup Q^5_{4,5} \cup Q^5_{5,2} \cup Q^5_{5,3} \cup Q^5_{5,4} \cup Q^5_{5,5}$.

The density is determined by

$$c^5_5(u,v) = \begin{cases} 0 & \textit{if} \quad (u,v) \in \mathbf{I}^2 \setminus (Q^5_{1,5} \cup Q^5_{2,3} \cup Q^5_{3,4} \cup Q^5_{4,2} \cup Q^5_{5,1}) \\ 5 & \textit{if} \quad (u,v) \in Q^5_{1,5} \cup Q^5_{2,3} \cup Q^5_{3,4} \cup Q^5_{4,2} \cup Q^5_{5,1}. \end{cases}$$

We notice that, in the same way that in the previous cases, the density of the copula is constant in every 2-box and its given by $s_{i,j}/\lambda^2(Q^5_{i,j})$*, for* $i, j \in I_m$*.*

We will see in the next theorem that the partition defined in 2.2.4 does not depend of the sample and we will see that the density of the copula is constant on every d-box:

Theorem 2.2.7. *Let* $2 \leq m \leq n$ *and let* $\underline{X}_1, ..., \underline{X}_n$ *be a random sample of size* n *from a random vector* $\underline{X}$ *of dimension* d*, with continuous joint distribution* H *or* d*-copula* C*, where* $\underline{X}_i = (X_{i,1}, ..., X_{i,d}) \in \mathbb{R}^d$*, for every* $i = 1, ..., n$*. Let* $U_n = \{\underline{Y}_1, ..., \underline{Y}_n\}$ *be the corresponding modified sample.*

Let $2 \leq m \leq n$ *fixed and define* $(R^m_{\underline{i}})_{\underline{i} \in (I_m)^d}$ *the uniform partition of size* m *of* $\mathbf{I}^d$ *as in equation* (2.8)*,* $s^{n,(m)}_{i_1,...,i_d}$ *as in equation* (2.9)*, the generalized transformation matrix* S^n_m *as in equation* (2.10)*, the partition* $(Q^m_{\underline{i}})_{\underline{i} \in (I_m)^d}$ *of* $\mathbf{I}^d$ *induced by* S^n_m *given in equation* (2.7) *, and* C^n_m *the sample copula of order* m *as in equation* (2.11)*. Then*

i) For the partitions of $(Q^m_{\underline{i}})_{\underline{i} \in (I_m)^d}$ *we know that* $0 = q_{1,0} < q_{1,1} < \cdots < q_{1,m} = 1$*, but we also have that*

$$q_{j,0} = q_{1,0} = 0, q_{j,1} = q_{1,1}, q_{j,2} = q_{1,2}, ..., q_{j,m} = q_{1,m} = 1 \textit{ for every } j \in \{2, 3, ..., d\}, \tag{2.12}$$

that is, in the d *coordinates the partition of* $\mathbf{I}$ *does not change. Even more, with probability one, the partition* $0 = q_{1,0} < q_{1,1} < \cdots < q_{1,m} = 1$ *only depends on* n *and* m*, and does not depend on the sample; in fact we have that*

$$q_{1,j} = \frac{1}{n} \cdot \left\lfloor \frac{j \cdot n}{m} \right\rfloor \textit{ for every } j \in \{0, 1, 2, ..., m\}, \tag{2.13}$$

where $\lfloor a \rfloor$ *denotes the greatest integer less than or equal to* a*.*

ii) Let λ^d *be the Lebesgue measure on the measurable space* $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$*, where* $\mathcal{B}(\mathbb{R}^d)$ *denotes the* σ*-algebra of Borel. If* C^n_m *denotes the sample copula of order* m*, let us denote by* c^n_m *its joint density function. Then*

$$c^n_m(u_1, ..., u_d) = s^{n,(m)}_{i_1,...,i_d}/\lambda^d(Q^m_{i_1,...,i_d}) \textit{ for every } (u_1, ..., u_d) \in Q^m_{i_1,...,i_d} \textit{ and } (i_1, ..., i_d) \in (I_m)^d. \tag{2.14}$$

Hence, the density is constant on every d-box $Q^m_{i_1,...,i_d}$ of the partition of $\mathbf{I}^d$ induced by S^n_m. Besides, if $M_d > n$ then exists at least one d-box $Q^m_{i_1,...,i_d}$ on which the density is zero. In fact, at most there are n d-boxes with positive density.
iii) For every $2 \leq m \leq n$, C^n_m is always a d-copula.
iv) Assume that m divides n, the the partition $(Q^m_{\underline{i}})_{\underline{i}\in(I_m)^d}$ of $\mathbf{I}^d$ coincides with the uniform partition $(R^m_{\underline{i}})_{\underline{i}\in(I_m)^d}$ of size m.
v) If $m = n$ there are exactly n elements of the partition $(Q^m_{\underline{i}})_{\underline{i}\in(I_m)^d} = (R^m_{\underline{i}})_{\underline{i}\in(I_m)^d}$ on which the density equals n^{d-1} and the remaining elements have density zero.

Proof: i) We observe that for every $k \in I_n$ and for every sample of size n, $\underline{Y}_k$ always has the form

$$\underline{Y}_k = \left(\frac{P_1(k)}{n}, \frac{P_2(k)}{n}, ..., \frac{P_d(k)}{n}\right), \tag{2.15}$$

where P_i is a permutation of I_n, for every $i \in I_d$.
We define, for every $k \in I_d$ and for every $l \in I_m$,

$$N_l(k) = \sum_{i=1}^{n} \mathbf{1}_{(\frac{l-1}{m},\frac{l}{m}]}\left(\frac{P_k(i)}{n}\right). \tag{2.16}$$

Since P_k is a permutation, for every $k \in I_d$, we have that $N_l(1) = N_l(2) = \cdots = N_l(d)$, for every $l \in I_m$. From this observation it follows that for every $j \in I_m$

$$q_{j,0} = q_{1,0} = 0, q_{j,1} = q_{1,1}, q_{j,2} = q_{1,2}, ..., q_{j,m} = q_{1,m} = 1, \tag{2.17}$$

and besides, for every $j \in \{0, 1, 2, ..., m\}$,

$$\begin{aligned} q_{1,j} &= \frac{1}{n}\sum_{l=1}^{j} N_l(1) \\ &= \frac{1}{n}\sum_{l=1}^{j}\sum_{i=1}^{n} \mathbf{1}_{\{\frac{l-1}{m}<\frac{P_1(i)}{n}\leq\frac{l}{m}\}} \\ &= \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{0<\frac{P_1(i)}{n}\leq\frac{j}{m}\}} \\ &= \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{0<P_1(i)\leq\frac{jn}{m}\}} \\ &= \frac{1}{n}\left\lfloor\frac{jn}{m}\right\rfloor. \end{aligned} \tag{2.18}$$

ii) Let $\mu_{C_m^n}$ be probability measure generated by C_m^n in $(\mathbf{I}^d, \mathcal{B}(\mathbf{I}^d))$. We notice that, by definition, for every $(i_1, ..., i_d) \in (I_m)^d$, we have that

$$\mu_{C_m^n}(Q^m_{i_1,...,i_d}) = s^{n,(m)}_{i_1,...,i_d}. \tag{2.19}$$

On the other hand, let c_m^n be the density function associated with C_m^n. Then for every $(i_1, ..., i_d) \in (I_m)^d$ we have that

$$\mu_{C_m^n}(Q^m_{i_1,...,i_d}) = \int_{q_{1,(i_d-1)}}^{q_{1,i_d}} \dots \int_{q_{1,(i_1-1)}}^{q_{1,i_1}} c_m^n(u_1, ..., u_d) du_1 \dots du_d, \tag{2.20}$$

and moreover

$$\lambda^d(Q^m_{i_1,...,i_d}) = \int_{q_{1,(i_d-1)}}^{q_{1,i_d}} \dots \int_{q_{1,(i_1-1)}}^{q_{1,i_1}} 1 du_1 \dots du_d = \prod_{k=1}^{d} (q_{1,i_k} - q_{1,(i_k-1)}). \tag{2.21}$$

By the definition of the sample d-copula of order m it follows that c_m^n is constant on every d-box. Hence

$$c_m^n(u_1, ..., u_d) = s^{n,(m)}_{i_1,...,i_d} / \lambda^d(Q^m_{i_1,...,i_d}) \text{ for every } (u_1, ..., u_d) \in Q^m_{i_1,...,i_d} \text{ and } (i_1, ..., i_d) \in (I_m)^d.$$

Now, we assume that $M_d > n$. We have that for $(u_1, ..., u_d) \in Q^m_{i_1,...,i_d}$, $c_m^n(u_1, ..., u_d) > 0$ if and only if $s^{n,(m)}_{i_1,...,i_d} > 0$. Since S_m^n is a generalized transformation matrix we can conclude that there are at most n d-boxes with positive density.

iii) We have already proved that C_m^n is a d-subcopula in $\{q_{1,0}, q_{1,1}, ..., q_{1,m}\}^d$. Then, by the proof of Lemma 2.3.5 in Nelsen's book [18], we have that C_m^n is a d-copula.

iv) We assume that m divides n. There exists $l \in \mathbb{N}$ such that $n = l \cdot m$. Thus, for every $j \in \{0, 1, 2, ..., m\}$ we have that

$$q_{1,j} = \frac{1}{n} \cdot \left\lfloor \frac{j \cdot l \cdot m}{m} \right\rfloor = \frac{1}{n} \cdot j \cdot l = \frac{j}{m}.$$

v) Let $m = n$. We notice that $N_1 = N_2 = \cdots = N_m = 1$, then there are exactly n elements of the partition $(Q^m_{\underline{i}})_{\underline{i} \in (I_m)^d}$ on which the density is positive and has the form

$$\frac{\frac{1}{n}}{\frac{1}{n^d}} = n^{d-1},$$

and the remaining elements have density zero.

The following definition can be understood as the maximum distance or the "distortion" between the uniform partition and the partition induced by the generalized transformation matrix given in 2.2.4.

Definition 2.2.8. *Let $2 \leq m \leq n$. Let, for every $k \in I_d$, $0 = r_{k,0} < 1/m = r_{k,1} < 2/m = r_{k,2} < \cdots < (m-1)/m = r_{k,m-1} < 1 = r_{k,m}$. Then $(r_{k,j})_{k \in I_d, j \in \{0,1,...,m\}}$ generates the partition induced by the uniform partition of size m. We define the distance between $(R^m_{\underline{i}})_{\underline{i} \in (I_m)^d}$ and $(Q^m_{\underline{i}})_{\underline{i} \in (I_m)^d}$ by*

$$e_m((R^m_{\underline{i}}), (Q^m_{\underline{i}})) = \max_{j \in \{0,1,...,m\}} |r_{1,j} - q_{1,j}|. \tag{2.22}$$

In the next proposition we establish a bound for the distance e_m defined above:

Proposition 2.2.9. *Let $2 \leq m \leq n$, $(R^m_{\underline{i}})_{\underline{i} \in (I_m)^d}$ as in (2.8), $(Q^m_{\underline{i}})_{\underline{i} \in (I_m)^d}$ as in (2.7) and e_m as in (2.22). Then*

$$\max_{2 \leq m \leq n} e_m((R^m_{\underline{i}}), (Q^m_{\underline{i}})) < \frac{1}{n} \tag{2.23}$$

Proof: Since $x - \lfloor x \rfloor < 1$ for every $x \in \mathbb{R}$, we notice that for every $j \in \{0, 1, ..., m\}$ and for every $2 \leq m \leq n$

$$\left| r_{1,j} - q_{1,j} \right| = \left| \frac{j}{m} - \left\lfloor \frac{jn}{m} \right\rfloor \frac{1}{n} \right| = \frac{1}{n} \left| \frac{jn}{m} - \left\lfloor \frac{jn}{m} \right\rfloor \right| < \frac{1}{n},$$

and hence the result follows.

Remark 2.2.10. *If n is a multiple of m then $e_m((R^m_{\underline{i}}), (Q^m_{\underline{i}})) = 0$, and this follow directly from Theorem 2.2.7 part iv).*

The following lemma shows that, in some cases, the sample d-copula of order m coincides (in the supremum distance) with the real copula.

Lemma 2.2.11. *Let $d \geq 2$ be an integer and let $n \geq 4$ be an even integer. Then there exists $2 \leq m \leq n$, C a d-copula and a sample of size n from C, such that*

$$\sup_{(u_1,...,u_d) \in \mathbf{I}^d} |C^n_m(u_1, ..., u_d) - C(u_1, ..., C_d)| = 0. \tag{2.24}$$

Proof: Let c be a function such that

$$c(u_1, ..., u_d) = \begin{cases} 2^{d-1} & \text{if} \quad (u_1, ..., u_d) \in [0, 1/2]^d \cup (1/2, 1]^d; \\ 0 & \text{if} \quad (u_1, ..., u_d) \in \mathbf{I}^d \setminus [0, 1/2]^d \cup (1/2, 1]^d. \end{cases}$$

We will see that C is a d-copula with density function c. In order to see that, let $u_j = 0$, for some $1 \leq j \leq d$. Then

$$C(u_1, ..., u_{j-1}, 0, u_{j+1}, ..., u_d) = \int_0^{u_1} \cdots \int_0^{u_{j-1}} \int_0^{0} \int_0^{u_{j+1}} \cdots \int_0^{u_d} c(x_1, ..., x_d) dx_d \cdots dx_1 = 0.$$

Let $0 \leq u_1 \leq 1$. Then

$$C(u_1, 1, ..., 1) = \int_0^1 \cdots \int_0^1 \int_0^{u_1} c(x_1, ..., x_d) dx_1 \cdots dx_d. \tag{2.25}$$

If $0 \leq u_1 \leq \frac{1}{2}$, then (2.25) is equal to

$$\int_0^1 \cdots \int_0^1 \int_0^1 \int_0^{u_1} 2^{d-1} dx_1 \cdots dx_d = \frac{1}{2^{d-1}} 2^{d-1} u_1 = u_1.$$

If $1/2 \leq u_1 \leq 1$, then (2.25) is equal to

$$\int_0^1 \cdots \int_0^1 2^{d-1} dx_1 \cdots dx_d = \frac{1}{2} + \frac{2^{d-1}}{2^{d-1}} \left(u_1 - \frac{1}{2} \right) = u_1.$$

Finally, as c is a nonnegative function, then clearly C is d-increasing. Hence C is a d-copula. Now, let $m = 2$. We notice that, by definition, all the mass of the d-copula C is accumulated in $R^2_{1,1,...,1}$ and $R^2_{2,2,...,2}$. Let $\underline{X}_1, ..., \underline{X}_d$ be a random sample of size n from the copula C, and assume that exactly $n/2$ elements of the sample fall in the d-box $R^2_{1,1,...,1}$. Then obviously the remaining $n/2$ elements fall in the d-box $R^2_{2,2,...,2}$ a.s. We observe that the modified sample $U_n = \underline{Y}_1, ..., \underline{Y}_n$ satisfies the same conditions as the sample. Then $s^{n,(2)}_{1,1,..,1} = s^{n,(2)}_{2,2,...,2} = 1/2$, and the remaining $s_{i_1,...,i_d} = 0$. Besides, by the theorem the density of the d-copula C^n_m is

$$c^n_m(u_1, ..., u_d) = \begin{cases} 2^{d-1} & \text{if} \quad (u_1, ..., u_d) \in [0, 1/2]^d \cup (1/2, 1]^d; \\ 0 & \text{if} \quad (u_1, ..., u_d) \in \mathbf{I}^d \setminus [0, 1/2]^d \cup (1/2, 1]^d. \end{cases}$$

Hence (2.24) is satisfied.

2.3 A comparison between the sample copula and the empirical copula

It is very important to mention that the sample d-copula of order m is far easier to compute than the empirical copula, the Bernstein copulas or the beta empirical copulas, see [13] and

[20]. All three have been used to estimate the true copula C, but as shown in [11], in all of these cases we may obtain better approximations to the true copula C using the sample copula. In fact, when the sample size is not small or the dimension is slightly large, in many cases the empirical copula, the Bernstein copulas or the beta empirical copula are impossible to evaluate in a standard computer.

In the first result of this section we determine a bound for the supremum distance between Π and M for the case $d \geq 2$ and also we give a bound for the case $d = 2$ for the supremum distance between Π and W.

Proposition 2.3.1. *Let $d \geq 2$ and Π_d, M_d and W_d as in* (1.6)*,* (1.3)*, and* (1.4)*, respectively. Then*

$$\sup_{(u_1,...,u_d)\in \mathbf{I}^d} |\Pi_d(u_1,...,u_d) - M_d(u_1,...,u_d)| = \frac{d-1}{d^{d/(d-1)}}, \tag{2.26}$$

where the supremum is attained at $u_1 = u_2 = \cdots = u_d = (d^{d/(d-1)})^{-1}$.
Besides,

$$\sup_{(u,v)\in \mathbf{I}^2} |\Pi_2(u,v) - M_2(u,v)| = \sup_{(u,v)\in \mathbf{I}^2} |\Pi_2(u,v) - W_2(u,v)| = \frac{1}{4}, \tag{2.27}$$

where the supremum is attained in $u = v = 1/2$ in both cases.

Proof: Let $0 \leq u_{(1)} \leq u_{(2)} \leq \cdots \leq u_{(d)} \leq 1$ such that $P(u_k) = u_j$, for $1 \leq k \leq d$, $1 \leq j \leq d$, and P is a permutation of I_d. Then

$$\begin{aligned}
\sup_{0\leq u_{(1)}\leq u_{(2)}\leq\cdots\leq u_{(d)}\leq 1} |\Pi_d(u_1,...,u_d) - M_d(u_1,...,u_d)| &= \sup_{0\leq u_{(1)}\leq u_{(2)}\leq\cdots\leq u_{(d)}\leq 1} \left| \prod_{i=1}^{d} u_{(i)} - u_{(1)} \right| \\
&= \sup_{0\leq u_{(1)}\leq u_{(2)}\leq\cdots\leq u_{(d)}\leq 1} u_{(1)} \left| \prod_{i=2}^{d} u_{(i)} - 1 \right| \\
&= \sup_{0\leq u_{(1)}\leq u_{(2)}\leq\cdots\leq u_{(d)}\leq 1} u_{(1)} \left(1 - \prod_{i=2}^{d} u_{(i)}\right) \\
&= \max_{0\leq u_{(1)}\leq 1} u_{(1)}(1 - u_{(1)}^{d-1}).
\end{aligned} \tag{2.28}$$

Let $f(u) = u(1 - u^{d-1})$, for $0 \leq u \leq 1$. Then $f'(u) = 1 - du^{d-1}$, and it follows hat $f'(u) = 0$ if and only if $u = (d^{1/(d-1)})^{-1}$. We observe that $f''(u) = -d(d-1)u^{d-2} < 0$, and hence f reaches a maximum at $u = (d^{1/(d-1)})^{-1}$ and

$$f\left(\frac{1}{d^{1/(d-1)}}\right) = \frac{1}{d^{1/(d-1)}}\left(1 - \left(\frac{1}{d^{1/(d-1)}}\right)^{d-1}\right) = \frac{d-1}{d^{d/(d-1)}},$$

and (2.26) is satisfied.

Now, let $0 \leq \alpha \leq 1$. We have that

$$\begin{aligned}\sup_{(u,v)\in\mathbf{I}^2} |\Pi_2(u,v) - W_2(u,v)| &= \sup_{(u,v)\in\mathbf{I}^2} |uv - \max(u+v-1,0)| \\ &= \max\{\sup_{u+v<1} |uv - \max(u+v-1,0)|, \sup_{u+v\geq 1} |uv - \max(u+v-1,0)|\} \\ &= \max\{\sup_{u+v<1} uv, \sup_{u+v\geq 1} |uv - u - v + 1|\} \\ &= \max\{\sup_{0\leq u<1/2} u(\alpha - u), \sup_{1/2\leq u\leq 1} (u-1)^2\}.\end{aligned} \tag{2.29}$$

Let $f(u) = u(\alpha - u)$, for $0 \leq u < 1/2$. Then $f'(u) = \alpha - 2u$, and $f'(u) = 0$ if only if $u = \alpha/2$. Besides, $f''(u) = -2 < 0$, then f reaches a maximum at $u = \alpha/2$. We observe that $f(\alpha/2) = \alpha^2/4$. Letting $\alpha \uparrow 1$ we have

$$\sup_{0\leq u<1/2} u(\alpha - u) = \lim_{\alpha\uparrow 1} \frac{\alpha^2}{4} = \frac{1}{4}.$$

Let $g(u) = (u-1)^2$, for $1/2 \leq u < 1$. We have that $g'(u) = 2(u-1) < 0$, and then the function g is strictly decreasing, hence reaches a maximum at $u = 1/2$. Then

$$\sup_{1/2\leq u\leq 1} (u-1)^2 = g\left(\frac{1}{2}\right) = \frac{1}{4},$$

and the result follows.

We notice that if we define $h(d) = (d-1)/(d^{d/(d-1)})$, then it can be easily proved that

$$\lim_{d\to\infty} h(d) = \lim_{d\to\infty} \frac{d-1}{d^{d/(d-1)}} = 1. \tag{2.30}$$

Let $\mathscr{C}_d$ be the set of all d-copulas. We notice that if we define, for every C_1, $C_2 \in \mathscr{C}_d$,

$$d_{sup}(C_1, C_2) = \sup_{(u_1,...,u_d)\in\mathbf{I}^d} |C_1(u_1, ..., u_d) - C_2(u_1, ..., u_d)|,$$

then $(\mathscr{C}_d, d_{sup})$ is obviously a metric space. Besides, by (1.5) and (2.30) if $C_1, C_2 \in \mathscr{C}_d$ then

$$0 \leq d_{sup}(C_1, C_2) \leq 1.$$

Definition 2.3.2. *Let $2 \leq m \leq n$ and let $\underline{X}_1, ..., \underline{X}_n$ be a random sample of size n from a random vector $\underline{X}$ of dimension d, with joint distribution H and copula C. Let $U_n = \{\underline{Y}_1, ..., \underline{Y}_n\}$ be the corresponding modified sample. Let C_n be the empirical copula defined as in* (2.1)*, and let C_m^n be the the sample copula of order m defined as in equation* (2.11)*. We define*

$$d_{sup_n}(C_n, C) = \max\left(\sup_{(i_1,...,i_d)\in I_n^d} \left| C_n\left(\frac{i_1}{n}, ..., \frac{i_d}{n}\right) - C\left(\frac{i_1}{n}, ..., \frac{i_d}{n}\right)\right|, \frac{1}{n}\right), \tag{2.31}$$

and

$$d_{sup_{n,(m)}}(C_m^n, C) = \sup_{(i_1,...,i_d)\in I_n^d} \left| C_m^n\left(\frac{i_1}{n}, ..., \frac{i_d}{n}\right) - C\left(\frac{i_1}{n}, ..., \frac{i_d}{n}\right)\right|. \tag{2.32}$$

Remark 2.3.3. *The function d_{sup_n} is never a metric, and $(\mathscr{C}_d, d_{sup_{n,(m)}})$ is a pseudometric space.*

The idea behind the definition in equation (2.31) arises from the Lemma 2.1.9, i.e., $d_{sup}(C_n, C) \geq 1/n$. Directly from the definition $d_{sup_n}(C_n, C) \geq 1/n$, and then d_{sup_n} is not a metric. We notice that from Theorem 2.2.7 C_m^n is a d-copula, and if C is a d-copula such that for every $(u_1, ..., u_d) \in \mathbf{I}^d$, $C_m^n(u_1, ..., u_d) = C(u_1, ..., u_d)$, then it is obvious that in particular $C_m^n(i_1/n, ..., i_d/n) = C(i_1/n, ..., i_d/n)$, for every $(i_1, ..., i_d) \in I_n^d$, and hence $d_{sup_{n,(m)}}(C_m^n, C) = 0$; but if $C_m^n(i_1/n, ..., i_d/n) = C(i_1/n, ..., i_d/n)$ for every $(i_1, ..., i_d) \in I_n^d$ does not imply that $C_m^n = C$. Hence $(\mathscr{C}_d, d_{sup_{n,(m)}})$ is a pseudometric space.

Since $\{(i_1/n, ..., i_d/n) | (i_1, ..., i_d) \in I_n^d\} \subset \mathbf{I}^d$, it is obvious that for every d-copula C

$$d_{sup_n}(C_n, C) \leq d_{sup}(C_n, C), \tag{2.33}$$

and

$$d_{sup_{n,(m)}}(C_m^n, C) \leq d_{sup}(C_m^n, C). \tag{2.34}$$

The following example shows that in (2.33) the inequality can be strict.

Example 2.3.4. *Let $d = 2$, $m = n = 2$ and $C = \Pi_2$. Let U_2 be the modified sample. We notice that U_2 only can be of the form $U_2 = \{(1/2, 1), (1, 1/2)\}$ or $U_2 = \{(1/2, 1/2), (1, 1)\}$; and takes every form with probability equal to 1/2. If the modified sample has the form $U_2 = \{(1/2, 1), (1, 1/2)\}$, we have that*

$$d_{sup_2}(C_2, \Pi_2) = \max\left(\left|0 - \frac{1}{4}\right|, \left|\frac{1}{2} - \frac{1}{2}\right|, \left|\frac{1}{2} - \frac{1}{2}\right|, \left|1 - 1\right|, \frac{1}{2}\right) = \frac{1}{2}.$$

Let $0 < \varepsilon < 1$. Then

$$d_{sup}(C_2, \Pi_2) = \sup_{(u,v)\in\mathbf{I}^2} |C_2(u, v) - \Pi_2(u, v)| \geq |C_2(1 - \varepsilon, 1 - \varepsilon) - (1 - \varepsilon)^2| = (1 - \varepsilon)^2.$$

Letting $\varepsilon \downarrow 0$, we have that

$$d_{sup}(C_2, \Pi_2) \geq \lim_{\varepsilon \downarrow 0} (1-\varepsilon)^2 = 1.$$

Hence

$$1 = d_{sup}(C_2, \Pi_2) > d_{sup_2}(C_2, \Pi_2) = \frac{1}{2}.$$

On the other hand, we have

$$d_{sup_{2,(2)}}(C_2^2, \Pi_2) = \left| C_2^2\left(\frac{1}{2}, \frac{1}{2}\right) - \Pi_2\left(\frac{1}{2}, \frac{1}{2}\right) \right| = \left| 0 - \frac{1}{4} \right| = \frac{1}{4},$$

and hence

$$1 = d_{sup}(C_2^2, \Pi_2) > d_{sup_{2,(2)}} = \frac{1}{4}.$$

Now, if the modified sample has the form $U_2 = \{(1/2, 1/2), (1, 1)\}$, we have that

$$d_{sup_2}(C_2, \Pi_2) = \max\left(\left| \frac{1}{2} - \frac{1}{4} \right|, \left| \frac{1}{2} - \frac{1}{2} \right|, \left| \frac{1}{2} - \frac{1}{2} \right|, \left| 1 - 1 \right|, \frac{1}{2} \right| \right) = \frac{1}{2},$$

and

$$\begin{aligned} d_{sup}(C_2, \Pi_2) &= \sup_{(u,v) \in \mathbf{I}^2} |C_2(u,v) - \Pi_2(u,v)| \\ &= \max(\sup_{(u,v) \in R^2_{1,1} \setminus \{(1/2),(1/2)\}} |0 - uv|, \sup_{(u,v) \in R^2_{1,2} \setminus \{(1/2,1/2)\}} |0 - uv|), \\ &\quad \sup_{(u,v) \in R^2_{2,1} \setminus \{(1/2,1/2)\}} |0 - uv|, |C_2(1/2,1/2) - \Pi_2(1/2,1/2)|, \sup_{(u,v) \in R^2_{2,2}} |1/2 - uv|) \\ &= \max\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{2}, \frac{1}{4}, \frac{1}{2} \right) = \frac{1}{2}. \end{aligned} \tag{2.35}$$

Hence

$$\frac{1}{2} = d_{sup}(C_2, \Pi_2) = d_{sup_2}(C_2, \Pi_2) = \frac{1}{2}.$$

On the other hand, we have

$$d_{sup_{2,(2)}}(C_2^2, \Pi_2) = \left| C_2^2\left(\frac{1}{2}, \frac{1}{2}\right) - \Pi_2\left(\frac{1}{2}, \frac{1}{2}\right) \right| = \left| \frac{1}{2} - \frac{1}{4} \right| = \frac{1}{4},$$

and hence

$$\frac{1}{2} = d_{sup}(C_2^2, \Pi_2) > d_{sup_{2,(2)}} = \frac{1}{4}.$$

The next example shows that in some cases $d_{sup_n}(C_n, C) = d_{sup}(C_n, C)$ on the grid $\{(i_1/n, ..., i_d/n) | (i_1, ..., i_d) \in \mathbf{I}^d\}$.

Example 2.3.5. *Let $\underline{X}_1, ..., \underline{X}_n$ be a random sample of size n from the copula M_d, defined as in* (1.3)*. Let $U_n = \{\underline{Y}_1, ..., \underline{Y}_n\}$ be the corresponding modified sample. Since M_d has as support the principal diagonal, we have that $\underline{X}_i = (u_i, u_i, ..., u_i)$, for all $i \in I_n$, with $u_i \in \mathbf{I}$. We can assume (reordering if necessary) that $u_1 < u_2 < \cdots < u_n$, and then $\underline{Y}_i = (i/n, i/n, ..., i/n)$, for all $i \in I_n$. Then, for $(v_1, ..., v_d) \in \mathbf{I}^d$ we have*

$$C_n(v_1, ..., v_d) = \begin{cases} 0 & \textit{if} \;\; \textit{there exists } i \in I_d \textit{ such that } v_i < 1/n \\ k/n & \textit{if} \;\; k/n \geq v_i \;\forall\; i \in I_d \textit{ and } \exists j \in I_d \textit{ such that } v_j < (k+1)/n, \end{cases}$$

where $k \in I_n$.
Let $(i_1, ..., i_d) \in I_n^d$. By definition

$$M_d\left(\frac{i_1}{n}, ..., \frac{i_d}{n}\right) = \frac{1}{n}\min(i_1, ..., i_d),$$

and hence

$$\sup_{(i_1,...,i_d) \in I_n^d} \left| C_n\left(\frac{i_1}{n}, ..., \frac{i_d}{n}\right) - M_d\left(\frac{i_1}{n}, ..., \frac{i_d}{n}\right) \right| = 0.$$

Definition 2.3.6. *Let μ and ν be two probabilities measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. We define the* ***total variation distance*** *between μ and ν, denoted by d_{TV}, as*

$$d_{TV}(\mu, \nu) = \sup_{A \in \mathcal{B}} |\mu(A) - \nu(A)|.$$

Recall that if f_μ and f_ν are the Radon-Nykodim's derivatives of μ and ν, respectively, then

$$0 \leq d_{TV}(\mu, \nu) = \frac{1}{2}\int_{\mathbb{R}^d} |f_\mu - f_\nu| d\lambda^d \leq 1,$$

where λ^d is the Lebesgue measure in $(\mathbb{R}^d, \mathcal{B})$.
Let $\underline{U}_1, ..., \underline{U}_n$ be a random sample from the product copula Π_2. Let $m = n$ and consider the d-sample copula of order m. In this case we are considering the uniform partition of size n, $(R^d_{\underline{i}})_{\underline{i} \in I_n^d}$, and according to Theorem 2.2.7, the density of the copula C_n^n is n^{d-1} in exactly n boxes and is zero in the remaining boxes. Let $J \subset I_n^d$ be the subset of n indices where the

density of C_n^n is positive and let . Let f_{Π_d} and $f_{C_n^n}$ be the densities of Π_d and C_n^d, respectively. Then

$$\begin{aligned} d_{TV}(\Pi_d, C_n^n) &= \frac{1}{2}\int_{\mathbf{I}^d} |f_{\Pi_d} - f_{C_n^d}| d\lambda^d \\ &= \frac{1}{2}\Bigg[\sum_{\underline{i}\in J}\int_{R^n_{\underline{i}}} |1-n^{d-1}|d\lambda^d + \sum_{\underline{i}\in(I^n_d\setminus J)}\int_{R^n_{\underline{i}}} |1-0|d\lambda^d\Bigg] \\ &= \frac{1}{2}\Bigg[\frac{n(n^{d-1}-1)}{n^d} + \frac{n^d-n}{n^d}\Bigg] \qquad (2.36) \\ &= \frac{1}{2}\Bigg[1-\frac{1}{n^{d-1}} + 1 - \frac{1}{n^{d-1}}\Bigg] \\ &= 1 - \frac{1}{n^{d-1}}. \end{aligned}$$

Hence

$$\lim_{n\to\infty} d_{TV}(\Pi_d, C_n^n) = 1,$$

and, moreover, for a fixed n we have

$$\lim_{d\to\infty} d_{TV}(\Pi_d, C_n^n) = 1.$$

2.4 Checkerboard approximation

In this section we define the Checkerboard approximation and give some basic results about it. Also, we present an important result on the convergence of this copula to the real copula.
The most important result of this section is the Glivenco-Cantelli's Theorem for the sample d-copula of order m. We use the checkerboard approximation to prove the theorem.
Using the notation in Lemma 2.3.5 in Nelsen's book [18], we give the followings definitions:

Definition 2.4.1. *Let $(a,b)\in\mathbf{I}^2$ and let C' be a 2-subcopula with finite domain $S_1\times S_2$. Let a_1 and a_2 be, respectively, the greatest and least elements of S_1 that satisfy $a_1\le a\le a_2$; and let b_1 and b_2 be, respectively, the greatest and least elements of S_2 that satisfy $b_1\le b\le b_2$. Clearly, if $a\in S_1$, then $a_1=a=a_2$, and if $b\in S_2$, then $b_1=b=b_2$. We define the quantities $\lambda_1(u,v)$ and $\mu_1(a,b)$ as follows*

$$\lambda_1(a,b)=\lambda_1=\begin{cases}(a-a_1)/(a_2-a_1) & \textit{if} \;\; a_1<a_2\\ 1 & \textit{if} \;\; a_1=a_2,\end{cases}$$

and

$$\mu_1(a,b) = \mu_1 = \begin{cases} (b-b_1)/(b_2-b_1) & \textit{if} \ \ b_1 < b_2 \\ 1 & \textit{if} \ \ b_1 = b_2. \end{cases}$$

By the proof of Lemma 2.3.5 in Nelsen's book [18], we have that, if we define

$$\begin{aligned} C(a,b) =&(1-\lambda_1(a,b))(1-\mu_1)(a,b)C'(a_1,b_1) + (1-\lambda_1(a,b))\mu_1(a,b)C'(a_1,b_2) \\ &+ \lambda_1(a,b)(1-\mu_1)(a,b)C'(a_2,b_1) + \lambda_1(a,b)\mu_1(a,b)C'(a_2,b_2), \end{aligned} \tag{2.37}$$

then C is a copula.

We give the following definition according to [15]:

Definition 2.4.2. *Let $\underline{X}$ be a random bivariate vector with joint distribution function H and copula C. Let $m \geq 1$. For every $(u,v) \in \mathbf{I}^2$, we define the* ***checkerboard approximation*** *of C using the uniform partition of size m, denoted as $C^{(m)}$, by*

$$\begin{aligned} C^{(m)}(u,v) =&\sum_{j=1}^{m}\sum_{i=1}^{m}\Bigg[1_{\left(\frac{i-1}{m},\frac{i}{m}\right]\times\left(\frac{j-1}{m},\frac{j}{m}\right]}(u,v)\Bigg((1-\lambda_1(u,v))(1-\mu_1(u,v))C\bigg(\frac{i-1}{m},\frac{j-1}{m}\bigg) \\ &+ (1-\lambda_1(u,v))\mu_1(u,v)C\bigg(\frac{i-1}{m},\frac{j}{m}\bigg) + \lambda_1(u,v)(1-\mu_1(u,v))C\bigg(\frac{i}{m},\frac{j-1}{m}\bigg) \\ &+ \lambda_1(u,v)\mu_1(u,v)C\bigg(\frac{i}{m},\frac{j}{m}\bigg)\Bigg]. \end{aligned} \tag{2.38}$$

Equation (2.38) is the same as in Lemma 2.3.5 in the proof of Sklar's Theorem in Nelsen's book using the copula C, i.e., $C^{(m)}$ assigns the mass uniformly in every box $R_{i,j}$. The checkerboard of order m, is an approximation of the density of a true d-copula C, based on a uniform partition of $\mathbf{I} = [0,1]$, given by $\{0, 1/m, 2/m, \ldots, (m-1)/m, 1\}$.
The following lemma establishes a Gilvenko-Cantelli's Theorem for the supremum distance between C^m and C:

Lemma 2.4.3. *Let $\underline{X}$ be a random bivariate vector with copula C. Let $m \geq 1$ and let $C^{(m)}$ be the checkerboard approximation as in* (2.38)*. Then, for every $m \geq 1$,*

$$\sup_{(u,v)\in\mathbf{I}^2} |C^{(m)}(u,v) - C(u,v)| < \frac{2}{m}.$$

In fact we have, in general

$$\sup_{(u_1,\ldots,u_d)\in\mathbf{I}^d} |C^{(m)}(u_1,\ldots,u_d) - C(u_1,\ldots,u_d)| \leq \frac{d}{2m}.$$

Moreover, it is well known, see [4], [5], [14] and [16], that $C^{(m)}$ is a good approximation of the true copula C even for moderate values of m. In fact, $C^{(m)}$ is a good density approximation for the true copula C. It is also trivial to see that $C^{(m)}$ has a density given by

$$c^{(m)}(u_1,\ldots,u_d) = V_C(\overline{R^m_{i_1,\ldots,i_d}})/\lambda^d(\overline{R^m_{i_1,\ldots,i_d}}) = M_d \cdot V_C(\overline{R^m_{i_1,\ldots,i_d}}) \text{ for every } (u_1,\ldots,u_d) \in R^m_{i_1,\ldots,i_d}, \tag{2.39}$$

where λ^d is the Lebesgue measure on the Borel space $(\mathbf{I}^d, \mathcal{B}(\mathbf{I}^d))$. Hence, the density is constant on each of the d-boxes of the uniform partition for every $(i_1,\ldots i_d) \in I^d_m$.

We notice that if we take $C = \Pi_2$ in Definition 2.4.2 then we have that, for every $m \geq 1$, $C^{(m)} = \Pi_2$. In order to see that let (a,b),a_1,a_2,b_1,b_2, C', λ_1 and μ_1 as in Definition 2.4.1. And let $C' = \Pi_2$, then by equation (2.37) we have

$$\begin{aligned}
C(a,b) &= \left(\frac{a_2-a}{a_2-a_1}\right)\left(\frac{b_2-b}{b_2-b_1}\right)a_1b_1 + \left(\frac{a_2-a}{a_2-a_1}\right)\left(\frac{b-b_1}{b_2-b_1}\right)a_1b_2 \\
&+ \left(\frac{a-a_1}{a_2-a_1}\right)\left(\frac{b_2-b}{b_2-b_1}\right)a_2b_1 + \left(\frac{a-a_1}{a_2-a_1}\right)\left(\frac{b-b_1}{b_2-b_1}\right)a_2b_2 \\
&= \frac{a_1a_2b_1b_2 - a_1a_2b_1b - a_1b_1b_2a + a_1b_1ab + a_1a_2b_2b - a_1a_2b_1b_2 - a_1b_2ab + a_1b_1b_2a}{(a_2-a_1)(b_2-b_1)} \\
&+ \frac{a_2b_1b_2a - a_2b_1ab - a_1a_2b_1b_2 + a_1a_2b_1b + a_2b_2ab - a_2b_1b_2a - a_1a_2b_2b + a_1a_2b_1b_2}{(a_2-a_1)(b_2-b_1)} \\
&= \frac{(a_1b_1 - a_1b_2 - a_2b_1 + a_2b_2)ab}{(a_2-a_1)(b_2-b_1)} = ab = \Pi_2(a,b).
\end{aligned} \tag{2.40}$$

Let $(u,v) \in \mathbf{I}^2$ and take $C = \Pi_2$ in equation (2.38). Then there exists a unique $(i,j) \in I_m$ such that $(u,v) \in ((i-1)/m, i/m] \times ((j-1)/m, j/m]$. Then using equation (2.40) we have that $C^{(m)} = \Pi_2$.

The next theorem shows that if we take C_n, the empirical copula, as the subcopula used in the proof of the Sklar's Theorem given in [18], then the resulting copula is C^n_m, the sample copula of order m.

Theorem 2.4.4. *Let $2 \leq m$ an let n be a multiple of m. Let $\underline{X}_1, ..., \underline{X}_n$ be a random sample of size n from a random vector $\underline{X}$ of dimension 2, with copula C;and let $U_n = \{\underline{Y}_1, ..., \underline{Y}_n\}$ be the corresponding modified sample. Let C_n be the empirical copula defined as in (2.1), and let C_m^n be the the sample copula of order m defined as in equation (2.11). Then for every $(u, v) \in \mathbf{I}^2$*

$$\begin{aligned} C_m^n(u,v) = & \sum_{j=1}^{m}\sum_{i=1}^{m}\Bigg[1_{\left(\frac{i-1}{m},\frac{i}{m}\right]\times\left(\frac{j-1}{m},\frac{j}{m}\right]}(u,v)\Bigg((1-\lambda_1(u,v))(1-\mu_1(u,v))C_n\left(\frac{i-1}{m},\frac{j-1}{m}\right) \\ & + (1-\lambda_1(u,v))\mu_1(u,v)C_n\left(\frac{i-1}{m},\frac{j}{m}\right) + \lambda_1(u,v)(1-\mu_1(u,v))C_n\left(\frac{i}{m},\frac{j-1}{m}\right) \\ & + \lambda_1(u,v)\mu_1(u,v)C_n\left(\frac{i}{m},\frac{j}{m}\right)\Bigg]. \end{aligned} \tag{2.41}$$

Proof: Let $(u, v) \in \mathbf{I}^2$, let $2 \leq m$ and let n be a multiple of m. Then there exists $(i, j) \in I_m^2$ such that $(u, v) \in R_{i,j}^2 = \left\langle \frac{i-1}{m}, \frac{i}{m}\right] \times \left\langle \frac{j-1}{m}, \frac{j}{m}\right]$, and besides there exists $k \in \mathbb{N}$ such that $n = mk$. Then

$$R_{i,j}^2 = \left\langle \frac{(i-1)n}{k}, \frac{ik}{n}\right] \times \left\langle \frac{(j-1)k}{n}, \frac{jk}{n}\right].$$

By definition of the 2-volume of a subcopula, we have that

$$\begin{aligned} V_{C_n}(R_{i,j}^2) &= C_n\left(\frac{ik}{n},\frac{jk}{n}\right) - C_n\left(\frac{(i-1)k}{n},\frac{jk}{n}\right) - C_n\left(\frac{ik}{n},\frac{(j-1)k}{n}\right) + C_n\left(\frac{(i-1)k}{n},\frac{(j-1)k}{n}\right) \\ &= \frac{1}{n}\Bigg[\sum_{l=1}^{n} 1_{\{Y_{l,1}\leq(ik)/n, Y_{l,2}\leq(jk)/n\}} - \sum_{l=1}^{n} 1_{\{Y_{l,1}\leq((i-1)k)/n, Y_{l,2}\leq(jk)/n\}} \\ &\quad - \sum_{l=1}^{n} 1_{\{Y_{l,1}\leq(ik)/n, Y_{l,2}\leq((j-1)k)/n\}} + \sum_{l=1}^{n} 1_{\{Y_{l,1}\leq((i-1)k)/n, Y_{l,2}\leq((j-1)k)/n\}}\Bigg] \\ &= \frac{1}{n}\sum_{l=1}^{n} 1_{\{((i-1)k)/n<Y_{l,1}\leq(ik)/n, ((j-1)k)/n<Y_{l,2}\leq(jk)/n\}} = \frac{card(R_{i,j}^2 \cap U_n)}{n} = s_{i,j}^{n,(m)}, \end{aligned} \tag{2.42}$$

where $s_{i,j}^{n,(m)}$ is defined as in equation (2.9).

Now, by the definition of the sample copula of order m, we have

$$\begin{aligned} C_m^n(u,v) &= \sum_{i'<i,j'<j} s_{i',j'}^{n,(m)} + \frac{nu-(i-1)k}{k}\sum_{j'<j} s_{i,j'}^{n,(m)} + \frac{nv-(j-1)k}{k}\sum_{i'<i} s_{i',j}^{n,(m)} \\ &+ \frac{(nu-(i-1)k)(nv-(j-1)k)}{k^2} s_{i,j}^{n,(m)} \\ &= \sum_{i'<i,j'<j} s_{i',j'}^{n,(m)} + \lambda_1(u,v)\sum_{j'<j} s_{i,j'}^{n,(m)} + \mu_1(u,v)\sum_{i'<i} s_{i',j}^{n,(m)} + \lambda_1(u,v)\mu_1(u,v) s_{i,j}^{n,(m)}. \end{aligned} \tag{2.43}$$

On the other hand, if the right side of (2.41) is equal to α, then

$$\begin{aligned} \alpha &= C_n\left(\frac{(i-1)k}{n},\frac{(j-1)k}{n}\right) + \lambda_1(u,v)\left[C_n\left(\frac{ik}{n},\frac{(j-1)k}{n}\right) - C_n\left(\frac{(i-1)k}{n},\frac{(j-1)k}{n}\right)\right] \\ &+ \mu_1(u,v)\left[C_n\left(\frac{(i-1)k}{n},\frac{jk}{n}\right) - C_n\left(\frac{(i-1)k}{n},\frac{(j-1)k}{n}\right)\right] + \lambda_1(u,v)\mu_1(u,v)\left[C_n\left(\frac{ik}{n},\frac{jk}{n}\right)\right. \\ &\left. - C_n\left(\frac{(i-1)k}{n},\frac{jk}{n}\right) - C_n\left(\frac{ik}{n},\frac{(j-1)k}{n}\right) + C_n\left(\frac{(i-1)k}{n},\frac{(j-1)k}{n}\right)\right] \\ &= \sum_{i'<i,j'<j} s_{i',j'}^{n,(m)} + \lambda_1(u,v)\sum_{j'<j} s_{i,j'}^{n,(m)} + \mu_1(u,v)\sum_{i'<i} s_{i',j}^{n,(m)} + \lambda_1(u,v)\mu_1(u,v) s_{i,j}^{n,(m)}, \end{aligned} \tag{2.44}$$

and the result follows.
The folloging theorem is a version of the Glivenko-Cantelli's Theorem for the sample copula of order m.

Theorem 2.4.5. *(Glivenko-Cantelli) Let $m \geq 2$ and let n be a multiple of m. Let $\underline{X}_1, ..., \underline{X}_n$ be a random sample of size n from a random vector $\underline{X}$ of dimension d, with copula C. Let C_m^n be the sample d-copula, defined as in equation (2.11). Let $\varepsilon > 0$; then there exists $N_\epsilon \in \mathbb{N}$ such that if $n, m \geq N_\varepsilon$, with $N_\varepsilon \leq m \leq n$, then*

$$\sup_{(u_1,...,u_d)\in\mathbf{I}^d} |C_m^n(u_1,...,u_d) - C(u_1,...,u_d)| < \varepsilon \ a.s \ [\mathbb{P}_C]. \tag{2.45}$$

Proof: We will do the proof for the case $d = 2$. Let $\varepsilon > 0$ and let C_n be the empirical copula as in (2.1). From the Glivenko-Cantelli's Theorem for the empirical copula, there exists $N_{1,\varepsilon} \in \mathbb{N}$ such that if $n \geq N_{1,\varepsilon}$, then

$$\sup_{(u,v)\in\mathbf{I}^2} |C_n(u,v) - C(u,v)| < \frac{\varepsilon}{2} \text{ a.s } [\mathbb{P}_C],$$

and it follows that, for every $(u,v) \in \mathbf{I}^2$,

$$|C_n(u,v) - C(u,v)| < \frac{\varepsilon}{2} \text{ a.s } [\mathbb{P}_C]. \tag{2.46}$$

Let $(u,v) \in \mathbf{I}^2$. Then there exist $i, j \in I_m$ such that $(u,v) \in R^m_{i,j} = \left\langle \frac{i-1}{m}, \frac{i}{m} \right] \times \left\langle \frac{j-1}{m}, \frac{j}{m} \right]$. Let $C^{(m)}$ be the checkerboard approximation, defined as in (2.38); let λ_1 and μ_1 as in Definition 2.4.1. Using Theorem 2.4.4, the triangle inequality and (2.46), we have

$$\begin{aligned}
|C^n_m(u,v) - C^{(m)}(u,v)| \leq & (1-\lambda_1(u,v))(1-\mu_1(u,v))\left|C_n\left(\frac{i-1}{m}, \frac{j-1}{m}\right) - C\left(\frac{i-1}{m}, \frac{j-1}{m}\right)\right| \\
& + (1-\lambda_1(u,v))\mu_1(u,v)\left|C_n\left(\frac{i-1}{m}, \frac{j}{m}\right) - C\left(\frac{i-1}{m}, \frac{j}{m}\right)\right| \\
& + \lambda_1(u,v)(1-\mu_1(u,v))\left|C_n\left(\frac{i}{m}, \frac{j-1}{m}\right) - C\left(\frac{i}{m}, \frac{j-1}{m}\right)\right| \\
& + \lambda_1(u,v)\mu_1(u,v)\left|C_n\left(\frac{i}{m}, \frac{j}{m}\right) - C\left(\frac{i}{m}, \frac{j}{m}\right)\right| \\
& < \Big[(1-\lambda_1(u,v))(1-\mu_1(u,v)) + (1-\lambda_1(u,v))\mu_1(u,v) \\
& + \lambda_1(u,v)(1-\mu_1(u,v)) + \lambda_1(u,v)\mu_1(u,v)\Big]\frac{\varepsilon}{2} = \frac{\varepsilon}{2} \text{ a.s } [\mathbb{P}_C].
\end{aligned} \tag{2.47}$$

Now, by the Lemma 2.4.3, we have that for every $m \leq 1$

$$\sup_{(u,v)\in\mathbf{I}^2} |C^{(m)}(u,v) - C(u,v)| < \frac{2}{m},$$

and by the Archimedean property there exists $N_{2,\varepsilon}$ such that if $m \geq N_{2,\varepsilon}$, then

$$\sup_{(u,v)\in\mathbf{I}^2} |C^{(m)}(u,v) - C(u,v)| < \frac{2}{m} < \frac{\varepsilon}{2} \text{ a.s } [\mathbb{P}_C].$$

Let $N_\varepsilon = \max(N_{1,\varepsilon}, N_{2,\varepsilon})$. If $n, m \geq N_\varepsilon$, with $N_\varepsilon \leq m \leq n$, then

$$\begin{aligned}
\sup_{(u,v)\in\mathbf{I}^2} |C^n_m(u,v) - C(u,v)| \leq & \sup_{(u,v)\in\mathbf{I}^2} |C^n_m(u,v) - C^{(m)}(u,v)| + \sup_{(u,v)\in\mathbf{I}^2} |C^{(m)}(u,v) - C(u,v)| \\
& < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \text{ a.s } [\mathbb{P}_C].
\end{aligned} \tag{2.48}$$

Chapter 3

Sample distribution under independence

In this chapter we assume that the modified sample comes from the product copula. Our goal is to find the distribution of the sample frequencies in the boxes generated by the uniform partition when we are sampling from the product copula. We will see that this distribution is well known. Moreover, we will find some sample moments, like means, variances and covariances.

3.1 Preliminaries

Remark 3.1.1. *If we consider the grid of $\mathbf{I}^2$, generated by the uniform partition of size n, then there exist $n!$ different ways in which can be observe the modified sample if the sample size is equal to n. That is because we have n different possibilities in the region $[0,1/n]\times[0,1]$, $n-1$ different possibilities in the region $(1/n,2/n]\times[0,1]$, and so on.*

We use the notation P^n_k to refer the **k-permutations of n**, i.e. ,

$$P^n_k = \frac{n!}{(n-k)!}.$$

We will use the following result in several proofs of this chapter. This is a well know result in combinatorial theory.

Proposition 3.1.2. *(Generalized Vandermonde's identity) Let $1 \leq m \leq n$, with $m,n \in \mathbb{Z}^+$. Let $n = l_1 + \cdots + l_j$, with $l_i \in \mathbb{Z}^+$ for every $i = 1,...,j$; and let $m = k_1 + \cdots + k_j$, with $k_i \in \mathbb{Z}^+ \cup \{0\}$. Then*

$$\binom{n}{m} = \sum_{k_1+\cdots k_j} \binom{l_1}{k_1} \cdots \binom{l_j}{k_j}.$$

We recall that if X has a hypergeometric distribution with parameters: N the population size, m the class 1 size, with $m \leq N$, and k the sample size, then

$$\mathbb{P}[X = x] = \frac{\binom{m}{x}\binom{N-m}{k-x}}{\binom{N}{k}},$$

and

$$\mathbb{E}[X] = \frac{km}{N},\ VaR[X] = \frac{km}{N}\frac{(N-m)}{N}\frac{(N-k)}{N-1}. \tag{3.1}$$

X counts the number of observations in the class 1.

Definition 3.1.3. *Let $\underline{X}_1, ..., \underline{X}_n$ be a random sample of size n from a random vector $\underline{X}$ of dimension 2 and $\underline{Y}_1, ..., \underline{Y}_n$ be the corresponding modified sample.Let $2 \leq m \leq n$, and assume that m divides n; let $l = n/m$. For $i, j \in I_m$, let*

$$R_{i,j} = \left\langle \frac{i-1}{m}, \frac{i}{m} \right] \times \left\langle \frac{j-1}{m}, \frac{j}{m} \right],$$

and let $N_{i,j}$ be the random variable that indicates the number of observations from the modified sample falling in the region $R_{i,j}$.

As m divides n, we have that, for every $i, j \in I_m$,

$$R_{i,j} = \langle (i-1)/m, i/m] \times \langle (j-1)/m, j/m] = \langle ((i-1)l)/n, (il)/n] \times \langle ((j-1)l)/n, (jl)/n].$$

For example, we can see the region $R_{1,1}$ in the figure 3.1.

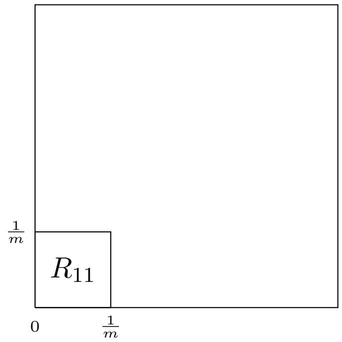


Figure 3.1: Region R_{11}

3.2 Distribution of the sample frequencies and moments

Lemma 3.2.1. *Let $n, m, l, R_{1,1}$ and $N_{1,1}$ as in the Definition 3.1.3. Then*

$$\mathbb{P}[N_{1,1} = n_{1,1}] = \frac{l!(n-l)!}{n!}\binom{l}{n_{1,1}}\binom{n-l}{l-n_{1,1}} = \frac{\binom{l}{n_{1,1}}\binom{n-l}{l-n_{1,1}}}{\binom{n}{l}},$$

i.e., the random variable $N_{1,1}$ has hypergeometric distributions with parameters: n the population size, l the class 1 size and l the sample size.

Proof: Let $A_1 = [0, l/n] \times (l/n, 1]$ and $A_2 = (l/n, 1] \times [0, 1]$. We notice that $N_{1,1}$ only can take values with positive probability in $\{0, 1, ..., l\}$. First select the number of ways in which we can select the $n_{1,1}$ observations for the coordinate X in the region $R_{1,1}$, that is $\binom{l}{n_{1,1}}$, as we can see in Figure 3.2; then we select the number of ways in which we can select the $n_{1,1}$ observations for the coordinate Y in the same region, that is $P^l_{n_{1,1}}$, as we can see in Figure 3.3. For the elements of the sample in the region A_1, we have $\binom{l-n_{1,1}}{l-n_{1,1}}$ ways to select the coordinate X and $P^{n-l}_{n-n_{1,1}}$ ways to select the coordinate Y. Finally, there are $(n-l)!$ ways to select the elements of the sample in the region A_2. Then

$$\begin{aligned}\mathbb{P}[N_{1,1} = n_{1,1}] &= \frac{\binom{l}{n_{1,1}} P^l_{n_{1,1}} \binom{l-n_{1,1}}{l-n_{1,1}} P^{n-l}_{l-n_{1,1}} (n-l)!}{n!} \\ &= \frac{\binom{l}{n_{1,1}} \frac{l!}{(l-n_{1,1})!} \frac{(n-l)!}{(n-2l+n_{1,1})!}(n-l)!}{n!} \\ &= \frac{\binom{l}{n_{1,1}}\binom{n-l}{l-n_{1,1}}(n-l)!l!}{n!} \\ &= \frac{\binom{l}{n_{1,1}}\binom{n-l}{l-n_{1,1}}}{\binom{n}{l}}.\end{aligned} \tag{3.2}$$

Theorem 3.2.2. *Let $n, m, l, R_{1,1}$ and $N_{1,1}$ as in the Definition 3.1.3. Then*

$$\mathbb{E}[N_{1,1}/n] = \frac{1}{m^2},\ \mathbb{E}[(N_{1,1}/n)^2] = \frac{l^2(l-1)^2}{n^3(n-1)} + \frac{1}{nm^2}, \tag{3.3}$$

and

$$Var[N_{1,1}/n] = \frac{l^2(l-1)}{n^3(n-1)} + \frac{1}{nm^2} - \left(\frac{1}{m^2}\right)^2.$$

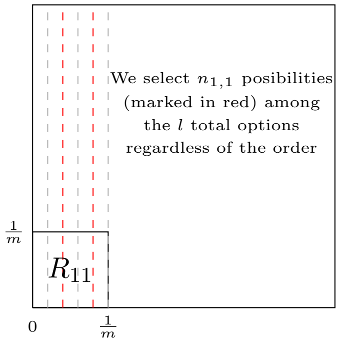


Figure 3.2:

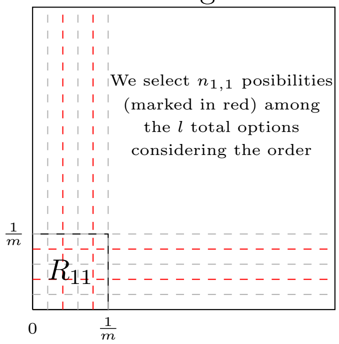


Figure 3.3:

Proof: According to the equations in (3.1), we have

$$\mathbb{E}[N_{1,1}/n] = \frac{1}{m^2},$$

and

$$\begin{aligned}
Var[N_{1,1}/n] &= \frac{l^2(n-l)^2}{n^4(n-1)} = \frac{l^2(n^2-2nl+l^2)}{n^4(n-1)} \\
&= \frac{l^2n^2 - 2nl^3 + l^4 + nl^4 - nl^4 + nl^2 - nl^2}{n^4(n-1)} \\
&= \frac{nl^2(l^2-2l+1) + l^2n(n-1) - l^4(n-1)}{n^4(n-1)} \\
&= \frac{l^2(l-1)^2}{n^3(n-1)} + \frac{l^2}{n^3} - \frac{l^4}{n^4} \\
&= \frac{l^2(l-1)}{n^3(n-1)} + \frac{1}{nm^2} - \left(\frac{1}{m^2}\right)^2.
\end{aligned} \tag{3.4}$$

Finally

$$\mathbb{E}[(N_{1,1}/n)^2] = Var[N_{1,1}/n] + (\mathbb{E}[N_{1,1}/n])^2 = \frac{l^2(l-1)}{n^3(n-1)} + \frac{1}{nm^2}.$$

Lemma 3.2.3. *Let $n, m, l, R_{i,j}$ and $N_{i,j}$ as in the Definition 3.1.3. Then for $r \in \{2, ..., m\}$*

$$\mathbb{P}[N_{1,1} = n_{1,1}, N_{1,r} = n_{1,r}] = \frac{l!(n-l)!}{n!}\binom{l}{n_{1,1}}\binom{l}{n_{1,r}}\binom{n-2l}{l-n_{1,1}-n_{1,r}}.$$

Proof: Let $A_1 = [0, l/n] \times ((l/n, 1] \backslash ((r-1)l/n, rl/n)))$, as in Figure 3.4, and $A_2 = (l/n, 1] \times [0, 1]$. We note that $0 \leq n_{1,1} \leq l$, $0 \leq n_{1,r} \leq l$ and are such that $0 \leq n_{1,1} + n_{1,r} \leq l$. Following the same idea as in the proof of the Lemma 3.2.1, first we select the number of ways in which we can select the $n_{1,1}$ observations for the coordinate X in the region $R_{1,1}$, that is to say, $\binom{l}{n_{1,1}}$; then there are $P^n_{n_{1,1}}$ ways in which we can select the coordinate Y for the same region. Now we select the number of ways in we which can select the $n_{1,r}$ for the coordinate X in the region $R_{1,r}$, that is $\binom{l-n_{1,1}}{n_{1,r}}$; then we select the number of ways for the coordinate Y in the same region, that is equal to $P^l_{n_{1,r}}$. Next, for the region A_1 we have $\binom{l-n_{1,1}-n_{1,r}}{l-n_{1,1}-n_{1,r}}$ ways in which can select the coordinate X and $P^{n-2l}_{l-n_{1,1}-n_{1,r}}$ ways for the coordinate Y. Finally, for the region A_2 we can select the sample in $(n-l)!$ ways. Hence

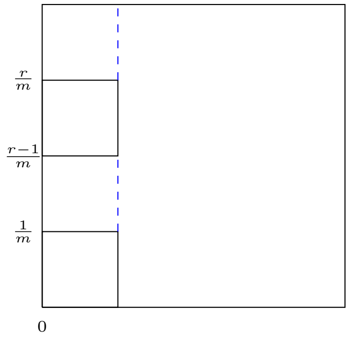


Figure 3.4: Region A_1

$$\mathbb{P}[N_{1,1} = n_{1,1}, N_{1,r} = n_{1,r}] = \frac{\binom{l}{n_{1,1}} P^l_{n_{1,1}} \binom{l-n_{1,1}}{n_{1,r}} P^l_{n_{1,r}} \binom{l-n_{1,1}-n_{1,r}}{l-n_{1,1}-n_{1,r}} P^{n-2l}_{l-n_{1,1}-n_{1,r}} (n-2l)!}{n!}$$
$$= \binom{l}{n_{1,1}} \frac{(n-l)!}{n!} \frac{l!}{(l-n_{1,1})!} \frac{(l-n_{1,1})!}{(l-n_{1,1}-n_{1,r})! n_{1,r}!} \frac{l!}{(l-n_{1,r})} \frac{(n-2l)!}{(n-3l+n_{1,1}+n_{1,r})!} \tag{3.5}$$
$$= \frac{l!(n-l)!}{n!}\binom{l}{n_{1,1}}\binom{l}{n_{1,r}}\binom{n-2l}{l-n_{1,1}-n_{1,r}}.$$

Using the Vandermonde's identity, we have that for every $r \in \{2, ..., l\}$

$$\begin{aligned}\sum_{0 \leq n_{1,1}+n_{1,r} \leq l} \mathbb{P}[N_{1,1} = n_{1,1}, N_{1,r} = n_{1,r}] &= \sum_{0 \leq n_{1,1}+n_{1,r} \leq l} \frac{l!(n-l)!}{n!} \binom{l}{n_{1,1}} \binom{l}{n_{1,r}} \binom{n-2l}{l-n_{1,1}-n_{1,r}} \\ &= \frac{l!(n-l)!}{n!} \sum_{0 \leq n_{1,1}+n_{1,r} \leq l} \binom{l}{n_{1,1}} \binom{l}{n_{1,r}} \binom{n-2l}{l-n_{1,1}-n_{1,r}} \\ &= \frac{l!(n-l)!}{n!} \binom{n}{l} \\ &= 1.\end{aligned} \tag{3.6}$$

Theorem 3.2.4. *Let $n, m, l, R_{i,j}$ and $N_{i,j}$ as in the Definition 3.1.3. Then for every $r \in \{2, ..., m\}$*

$$Cov[N_{1,1}/n, N_{1,r}/n] = \frac{l^3(l-1)}{n^3(n-1)} - \left(\frac{1}{m^2}\right)^2. \tag{3.7}$$

Proof:

$$\begin{aligned}\mathbb{E}[N_{1,1}N_{1,r}] &= \sum_{0 \leq n_{1,1}+n_{1,r} \leq l} n_{1,1}n_{1,r} \frac{l!(n-l)!}{n!} \binom{l}{n_{1,1}} \binom{l}{n_{1,r}} \binom{n-2l}{l-n_{1,1}-n_{1,r}} \\ &= \frac{l!(n-l)!}{n!} \sum_{0 \leq n_{1,1}+n_{1,r} \leq l} n_{1,1}n_{1,r} \frac{l!}{(l-n_{1,1})!n_{1,1}!} \frac{l!}{(l-n_{1,r})!n_{1,r}!} \binom{n-2l}{l-n_{1,1}-n_{1,r}} \\ &= \frac{l^2 l!(n-l)!}{n!} \sum_{0 \leq n_{1,1}+n_{1,r} \leq l} \binom{l-1}{n_{1,1}-1} \binom{l-1}{n_{1,r}-1} \binom{n-2l}{l-n_{1,1}-n_{1,r}}.\end{aligned} \tag{3.8}$$

Let $j_1 = n_{1,1} - 1$ and $j_r = n_{1,r}$. It follows that $j_1 + j_r = n_{1,1} + n_{1,r} - 2$, and hence $0 \leq j_1 + j_r \leq l - 2$. Then by the above and by the Vandermonde's identity

$$\begin{aligned}\mathbb{E}[N_{1,1}N_{1,r}] &= \frac{l^2 l!(n-l)!}{n!} \sum_{0 \leq j_1+j_r \leq l-2} \binom{l-1}{j_1} \binom{l-1}{j_r} \binom{n-2l}{(l-2)-(j_1+j_r)} \\ &= \frac{l^2 l!(n-l)!}{n!} \binom{n-2}{l-2} \\ &= \frac{l^2(n-l)!l!(n-2l)!}{n!(n-l)!(l-2)!} \\ &= \frac{l^3(l-1)}{n(n-1)}.\end{aligned} \tag{3.9}$$

Then

$$\mathbb{E}[(N_{1,1}/n)(N_{1,r}/n)] = \frac{l^3(l-1)}{n^3(n-1)},$$

and

$$Cov[N_{1,1}/n, N_{1,r}/n] = \frac{l^3(l-1)}{n^3(n-1)} - \left(\frac{1}{m^2}\right)^2.$$

Lemma 3.2.5. *Let $n, m, , l, R_{1,1}, R_{2,2}, N_{1,1}$ and $N_{2,2}$ as in the Definition 3.1.3. Then*

$$\mathbb{P}[N_{1,1} = n_{1,1}, N_{2,2} = n_{2,2}] = \frac{(l!)^2(n-2l)!}{n!} \sum_{x_1+x_2=l-n_{1,1}} \sum_{y_1+y_2=l-n_{2,2}} \left[\binom{l}{n_{1,1}}\binom{l}{x_1}\binom{n-2l}{x_2}\right.$$
$$\left.\binom{l-n_{1,1}}{y_1}\binom{l-x_1}{n_{2,2}}\binom{n-2l-x_2}{y_2}\right]. \tag{3.10}$$

Proof: Let $A_1 = (2l/n, 1] \times [0, l/n]$, $A_2 = (2l/n, 1] \times (l/n, 2l/n]$, and $A_3 = [0,1] \times (2l/n, 1]$. The number of ways in which we can select the $k_{1,1}$ observations for the coordinate X in the region $R_{1,1}$ is $\binom{l}{n_{1,1}}$ and is equal to $P^l_{n_{1,1}}$ for the coordinate Y. In the region $R_{2,1}$ we can select x_1 observations from the sample in $\binom{l}{x_1}$ ways for the coordinate X and $P^{l-n_{1,1}}_{x_1}$ ways for the coordinate Y; and for the region A_1, we can select x_2 observations for the coordinate X in $\binom{n-2l}{x_2}$ ways and the coordinate Y in $P^{l-n_{1,1}-x_1}_{x_2}$. We note that the condition $n_{1,1} + x_1 + x_2 = l$ must be satisfied. Now, for the region $R_{1,2}$ we can select y_1 observations in $\binom{l-n_{1,1}}{y_1}$ ways for the coordinate X and $P^l_{y_1}$ ways for the coordinate Y; the numbers of ways in which we can select $n_{2,2}$ observations in the region $R_{2,2}$ for the coordinate X is $\binom{l-x_1}{n_{2,2}}$, and for the coordinate Y is $P^{l-y_1}_{n_{2,2}}$; and for the region A_2 we can select y_2 observations in $\binom{n-2l-x_2}{y_2}$ different ways for the coordinate X and $P^{l-y_1-n_{2,2}}_{y_2}$. We note again that $y_1 + n_{2,2} + y_2 = l$ must be satisfied. Finally, for the region called A_3, we can select the observations in $(n-2l)!$ ways. We can see all this in the Figure 3.5. Then

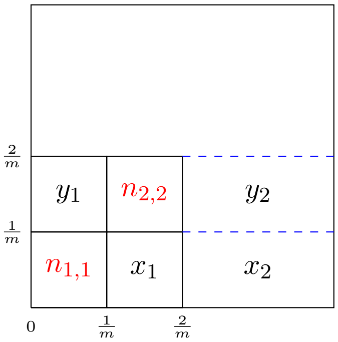


Figure 3.5:

$$\mathbb{P}[N_{1,1}=n_{1,1},N_{2,2}=n_{2,2}] = \sum_{x_1+x_2=l-n_{1,1}} \sum_{y_1+y_2=l-n_{2,2}} \Bigg[\binom{l}{n_{1,1}} P^{l}_{n_{1,1}} \binom{l}{x_1} P^{l-n_{1,1}}_{x_1} \binom{n-2l}{x_2} P^{l-n_{1,1}-x_1}_{x_2}$$

$$\binom{l-n_{1,1}}{y_1} P^{l}_{y_1} \binom{l-x_1}{n_{2,2}} P^{l-y_1}_{n_{2,2}} \binom{n-2l-x_2}{y_2} P^{l-y_1-n_{2,2}}_{y_2} \frac{(n-2l)!}{n!}\Bigg]$$

$$= \frac{(n-2l)!}{n!} \sum_{x_1+x_2=l-n_{1,1}} \sum_{y_1+y_2=l-n_{2,2}} \Bigg[\binom{l}{n_{1,1}} \frac{l!}{(l-n_{1,1})!} \binom{l}{x_1}$$

$$\frac{(l-n_{1,1})!}{(l-n_{1,1}-x_1)!} \binom{n-2l}{x_2} \frac{(l-n_{1,1}-x_1)!}{(l-n_{1,1}-x_1-x_2)!} \binom{l-n_{1,1}}{y_1} \frac{l!}{(l-y_1)!}$$

$$\binom{l-x_1}{n_{2,2}} \frac{(l-y_1)!}{(l-y_1-n_{2,2})!} \binom{n-2l-x_2}{y_2} \frac{(l-y_1-n_{2,2})!}{l-y_1-n_{2,2}-y_2)!}\Bigg]$$

$$= \frac{(l!)^2(n-2l)!}{n!} \sum_{x_1+x_2=l-n_{1,1}} \sum_{y_1+y_2=l-n_{2,2}} \Bigg[\binom{l}{n_{1,1}} \binom{l}{x_1} \binom{n-2l}{x_2}$$

$$\binom{l-n_{1,1}}{y_1} \binom{l-x_1}{n_{2,2}} \binom{n-2l-x_2}{y_2}\Bigg]$$

Using the Vandermonde's identity we have

$$\begin{aligned}
\sum_{n_{1,1}=0}^{l}\sum_{n_{2,2}=0}^{l}\mathbb{P}[N_{1,1}=n_{1,1},N_{2,2}=n_{2,2}] &= \sum_{n_{1,1}=0}^{l}\sum_{n_{2,2}=0}^{l}\frac{(l!)^2(n-2l)!}{n!}\sum_{x_1+x_2=l-n_{1,1}}\sum_{y_1+y_2=l-n_{2,2}} \\
&\left[\binom{l}{n_{1,1}}\binom{l}{x_1}\binom{n-2l}{x_2}\binom{l-n_{1,1}}{y_1}\binom{l-x_1}{n_{2,2}}\binom{n-2l-x_2}{y_2}\right] \\
&= \frac{(l!)^2(n-2l)!}{n!}\sum_{n_{1,1}=0}^{l}\sum_{x_1+x_2=l-n_{1,1}}\left[\binom{l}{n_{1,1}}\binom{l}{x_1}\binom{n-2l}{x_2}\right. \\
&\left.\sum_{n_{2,2}=0}^{l}\sum_{y_1+y_2=l-n_{2,2}}\binom{l-n_{1,1}}{y_1}\binom{l-x_1}{n_{2,2}}\binom{n-2l-x_2}{y_2}\right] \\
&= \frac{(l!)^2(n-2l)!}{n!}\sum_{n_{1,1}=0}^{l}\sum_{x_1+x_2=l-n_{1,1}}\left[\binom{l}{n_{1,1}}\binom{l}{x_1}\binom{n-2l}{x_2}\right. \\
&\left.\binom{n-n_{1,1}-x_1-x_2}{l}\right] \\
&= \frac{(l!)^2(n-2l)!}{n!}\binom{n}{l}\binom{n-l}{l} \\
&= \frac{(l!)^2(n-2l)!}{n!}\frac{n!}{(n-l)!l!}(n-l)!(n-2l)!l! \\
&= 1.
\end{aligned} \tag{3.11}$$

Theorem 3.2.6. *Let $n, m, , l, N_{1,1}$ and $N_{2,2}$ as in the Definition 3.1.3. Then*

$$Cov[N_{1,1}/n, N_{2,2}/n] = \frac{l^4}{n^3(n-1)} - \frac{1}{m^4}.$$

Proof:

$$\mathbb{E}[N_{1,1}N_{2,2}] = \sum_{n_{1,1}=0}^{l}\sum_{n_{2,2}=0}^{l}\frac{(l!)^2(n-2l)!}{n!}n_{1,1}n_{2,2}\sum_{x_1+x_2=l-n_{1,1}}\ \sum_{y_1+y_2=l-n_{2,2}}$$
$$\left[\binom{l}{n_{1,1}}\binom{l}{x_1}\binom{n-2l}{x_2}\binom{l-n_{1,1}}{y_1}\binom{l-x_1}{n_{2,2}}\binom{n-2l-x_2}{y_2}\right]$$

$$= \frac{(l!)^2(n-2l)!}{n!}\sum_{n_{1,1}=1}^{l}\sum_{x_1+x_2=l-n_{1,1}}\left[n_{1,1}\frac{l(l-1)!}{n_{1,1}(n_{1,1}-1)!(l-n_{1,1})!}\right. \tag{3.12}$$
$$\binom{l}{x_1}\binom{n-2l}{x_2}\sum_{n_{2,2}=1}^{l}\sum_{y_1+y_2=l-n_{2,2}} n_{2,2}\frac{(l-x_1)(l-x_1-1)!}{n_{2,2}(n_{2,2}-1)!(l-x_1-n_{2,2})!}$$
$$\left.\binom{l-n_{1,1}}{y_1}\binom{n-2l-x_2}{y_2}\right].$$

Let $j_1 = n_{1,1} - 1$ and $j_2 = n_{2,2} - 1$. Then

$$\mathbb{E}[N_{1,1}N_{2,2}] = \frac{(l!)^2(n-2l)!l}{n!}\sum_{j_1=0}^{l-1}\sum_{x_1+x_2=l-1-j_1}\left[\binom{l-1}{j_1}\binom{l}{x_1}\binom{n-2l}{x_2}\right.$$
$$\left.(l-x_1)\sum_{j_2=0}^{l-1}\sum_{y_1+y_2=l-1-j_2}\binom{l-x_1-1}{j_2}\binom{l-1-j_1}{y_1}\binom{n-2l-x_2}{y_2}\right], \tag{3.13}$$

and by the Vandermonde's identity

$$\mathbb{E}[N_{1,1}N_{2,2}] = \frac{(l!)^2(n-2l)!l}{n!}\sum_{j_1=0}^{l-1}\sum_{x_1+x_2=l-1-j_1}\left[\binom{l-1}{j_1}\frac{l(l-1)!}{(l-x_1)!x_1!}(l-x_1)\right.$$
$$\left.\binom{n-2l}{x_2}\binom{n-2l}{x_2}\right]$$

$$= \frac{(l!)^2(n-2l)!l^2}{n!}\sum_{j_1=0}^{l-1}\sum_{x_1+x_2=l-1-j_1}\binom{l-1}{j_1}\binom{l-1}{x_1}\binom{n-2l}{x_2}\binom{n-2l}{x_2}$$

$$= \frac{(l!)^2l^2(n-2l)!}{n!}\binom{n-2}{l-1}\binom{n-l-1}{l-1}$$

$$= \frac{l^4}{n(n-1)}.$$

Then

$$\mathbb{E}[(N_{1,1}/n)(N_{2,2}/n)] = \frac{l^4}{n^3(n-1)},$$

and hence

$$Cov[N_{1,1}/n, N_{2,2}/n] = \frac{l^4}{n^3(n-1)} - \frac{1}{m^4}.$$

Theorem 3.2.7. *Let $2 \leq m \leq n$, $n \in \mathbb{N}$, and assume that m divides n, i.e. $l = n/m$. Let $I_m = \{1, ..., m\}$, $(R_{i,j})_{(i,j)\in I_m}$ be the uniform partition of size m of $\mathbf{I}^2$ and $N_{i,j}$ be the random variable that indicate the number of observations falling in $R_{i,j}$, for every $i, j \in I_m$, when we consider the modified sample of size n from the product copula; for $i, j \in I_m$, let $n_{i,j}$ be zero or a positive integer, satisfying the following conditions*

$$\sum_{j=1}^{m} n_{i,j} = l \textit{ for all } \ i \in I_m,$$

and

$$\sum_{i=1}^{m} n_{i,j} = l \textit{ for all } \ j \in I_m.$$

Then

$$\mathbb{P}\Big[\bigcap_{i,j\in I_m} \{N_{i,j} = n_{i,j}\}\Big] = \frac{(l!)^{2m}}{n! \displaystyle\prod_{i,j\in I_m} n_{i,j}!}.$$

Proof: First we select $n_{1,1}$ elements from the sample in the region $R_{1,1}$, for the coordinate X we have $\binom{l}{n_{1,1}}$ ways and for the coordinate Y is given by $P^{l}_{n_{1,1}}$; after that, we select $n_{1,2}$ elements from the sample in the region $R_{1,2}$, this can be done in $\binom{l-n_{1,1}}{n_{2,2}}$ ways for the coordinate X and $P^{l}_{n_{2,2}}l$ ways for the coordinate Y; and so on, i.e., we select $n_{1,m}$ elements from the in the region $R_{1,m}$ in $\binom{l-\sum_{j=1}^{m-1} n_{1,j}}{n_{1,m}}$ ways for the coordinate X and $P^{l}_{n_{1,m}}$ ways for the coordinate Y. Now, for the region $R_{2,1}$ there are $\binom{l}{n_{2,1}}$ and $P^{l-n1,1}_{n_{2,1}}$ ways in which we can select $n_{1,2}$ elements from the sample for the coordinates X and Y, respectively; for the region $R_{2,2}$ we can select $n_{2,2}$ points from the sample in $\binom{l-n_{2,1}}{n_{2,2}}$ ways for the X axis and $P^{l-n_{1,2}}_{n_{2,2,}}$ ways for the Y axis; and so on, i.e., for the region $R_{2,m}$ the number of ways in which we can select $n_{2,m}$ elements for the sample is given by $\binom{l-\sum_{j=1}^{m-1} n_{2,j}}{n_{2,m}}$ for the coordinate X and $P^{l-n_{1,m}}_{n_{2,m}}$ for the coordinate Y.

$$
\begin{aligned}
\mathbb{P}\Bigg[\bigcap_{i,j\in I_m}\{N_{i,j}=n_{i,j}\}\Bigg] &= \binom{l}{n_{1,1}}P^{l}_{n_{1,1,}}\binom{l-n_{1,1}}{n_{1,2}}P^{l}_{n_{1,2}}\cdots\binom{l-\sum_{j=1}^{m-1}n_{1,j}}{n_{1,m}}P^{l}_{n_{1,m}}\\
&*\binom{l}{n_{2,1}}P^{l-n_{1,1}}_{n_{2,1,}}\binom{l-n_{2,1}}{n_{2,2}}P^{l-n_{1,2}}_{n_{2,2}}\cdots\binom{l-\sum_{j=1}^{m-1}n_{2,j}}{n_{2,m}}P^{l-n_{1,m}}_{n_{2,m}}\\
&\vdots\\
&*\binom{l}{n_{m,1}}P^{l-\sum_{i=1}^{m-1}n_{i,1}}_{n_{m,1,}}\binom{l-n_{m,1}}{n_{m,2}}P^{l-\sum_{i=1}^{m-1}n_{i,2}}_{n_{m,2}}\cdots\binom{l-\sum_{j=1}^{m-1}n_{m,j}}{n_{m,m}}P^{l-\sum_{i=1}^{m-1}n_{i,m}}_{n_{m,m}}\\
&=\frac{l!}{(l-n_{1,1})!n_{1,1}!}\frac{l!}{(l-n_{1,1})!}\frac{(l-n_{1,1})!}{(l-(n_{1,1}+n_{1,2}))!n_{1,2}!}\frac{l!}{(l-n_{1,2})!}\\
&\cdots\frac{(l-\sum_{j=1}^{m-1}n_{1,j})!}{(l-\sum_{j=1}^{m}n_{1,j})!n_{1,m}!}\frac{l!}{(l-n_{1,m})!}\\
&\frac{l!}{(l-n_{2,1})!n_{2,1}!}\frac{(l-n_{1,1})!}{(l-(n_{1,1}+n_{2,1}))!}\frac{(l-n_{2,1})!}{(l-(n_{2,1}+n_{2,2}))!n_{2,2}!}\frac{(l-n_{1,2})!}{(l-(n_{1,2}+n_{2,2}))!}\\
&\cdots\frac{(l-\sum_{j=1}^{m-1}n_{2,j})!}{(l-\sum_{j=1}^{m}n_{2,j})!n_{2,m}!}\frac{(l-n_{1,m})!}{(l-(n_{1,m}+n_{2,m}))!}\\
&\vdots\\
&\frac{l!}{(l-n_{m,1})!n_{m,1}!}\frac{(l-\sum_{i=1}^{m-1}n_{i,1})!}{(l-\sum_{i=1}^{m}n_{i,1})!}\frac{(l-n_{m,1})!}{(l-(n_{m,1}+n_{m,2}))!n_{m,2}!}\frac{(l-\sum_{i=1}^{m-1}n_{i,2})!}{(l-\sum_{i=1}^{m}n_{i,2})!}\\
&\cdots\frac{(l-\sum_{j=1}^{m-1}n_{m,j})!}{(l-(\sum_{j=1}^{m}n_{m,j}-n_{m,m}))!n_{m,m}!}\frac{(l-\sum_{i=1}^{m-1}n_{i,m})!}{(l-(\sum_{i=1}^{m}n_{i,m}-n_{m,m}))!}\\
&=\frac{(l!)^{2m}}{n!\prod_{i,j\in I_m}n_{i,j}!}.
\end{aligned}
$$

3.3 General case, $d \geq 2$

The objective in this section is to discuss how we can generalize some of the ideas presented above to the case $d \geq 2$, concerning the moments and the joint distribution.

In the same way as in the case when $d = 2$, let $2 \leq m \leq n$, such that m divides n.

Definition 3.3.1. *Let $d \geq 2$ and let let $2 \leq m \leq n$, such that m divides n, that is, there exists $l \in \mathbb{Z}$ such that $l = n/m$. Let $\underline{i} = (i, i, ..., i)$, for $i = 1, 2$, that is, $\underline{i}$ is a d-dimensional vector with every entry equal to 1. Let $N_{\underline{1}}$ be the random variable that indicates the number of observations in the d-box $R_{\underline{1}} = [0, l/n]^d$, and let $N_{\underline{2}}$ be the random variable that indicates the number of observations in the d-box $R_{\underline{2}} = (l/n, 2l/n]^d$.*

Theorem 3.3.2. *Let $d, n, m, l, N_{\underline{1}}, N_{\underline{2}}, R_{\underline{1}}$ and $R_{\underline{2}}$ as in Definition 3.3.1.Then*

i) $E(N_{\underline{1}}/n) = \frac{1}{m^d}$,

ii) $E((N_{\underline{1}}/n)^2) = \frac{l^d(l-1)^d}{n^{d+1}(n-1)^{d-1}} + \frac{1}{nm^d}$,

iii) $Var(N_{\underline{1}}/n) = \frac{l^d(l-1)^d}{n^{d+1}(n-1)^{d-1}} + \frac{1}{nm^d} - \frac{1}{m^{2d}}$,

iv) $Cov(N_{\underline{1}}/n, N_{\underline{2}}/n) = \frac{l^{2d}}{n^{d+1}(n-1)^{d-1}} - \frac{1}{m^{2d}}$.

Theorem 3.3.3. *Let $d \geq 2$, and let $N_{i_1 \cdots i_d}$ be the random variable that indicates the number of observations in the box $R_{i_1 \cdots i_d} = \langle (i_1 - 1)/m, i_1/m] \times \cdots \times \langle (i_d - 1)/m, i_d/m]$. Then*

$$\mathbb{P}\Bigg[\bigcap_{i_1, \cdots, i_d \in I_m} \{N_{i_1 \cdots i_d} = n_{i_1 \cdots i_d}\} \Bigg] = \frac{(l!)^{dm}}{(n!)^{d-1} \prod_{i_1, \cdots, i_d \in I_m} n_{i_1 \ldots i_d}!}.$$

For more details, see [10].

3.4 Additional comments

The results presented above are valid in general, but we are interested only in the cases $m = 2$ and $m = 3$.Then we note the following, if $d = 2$ we have:
If $m = 2$ then $N_{1,1}$ determines the distribution of C_2^n.
If $m = 3$ then $N_{1,1}, N_{1,2}, N_{2,1}$ and $N_{2,2}$ give the distribution of C_3^n.
In the following chapter we will give a characterization of independence using the distribution

of C_2^n given C_3^n. Using the present chapter, we could give the exact distribution of the statistics that we will propose, but we will not present that result in this work. Instead of that, we notice the following: because C_2^n and C_3^n are copulas and how $[0,1/3]^2 \subset [0,1/2]^2 \subset [0,2/3]^2$ we have

$$C_3^n\left(\frac{1}{3},\frac{1}{3}\right) \leq C_2^n\left(\frac{1}{2},\frac{1}{2}\right) \leq \min\left\{C_3^n\left(\frac{2}{3},\frac{2}{3}\right),\frac{1}{2}\right\},$$

and

$$\frac{1}{3} \leq C_3^n\left(\frac{2}{3}\right) \leq \frac{2}{3}.$$

Hence, C_2^n and C_3^n are not independent.

Chapter 4

A new test of independence

In this chapter we provide a very simple characterization of Π^d, the independence copula in dimension $d \geq 2$, in terms of the checkerboard approximations of order $m = 2$ and $m = 3$. As we will see in the third section, the independence copula Π^d, satisfies that $\Pi^d(u_1, \ldots, u_d) = C^{(m)}(u_1, \ldots, u_d)$ for every $(u_1, \ldots, u_d) \in \mathbf{I}^d$ and for every $m \geq 2$. However, for the converse we only need the equality for $m = 2$ and $m = 3$.
Let H_0 denote the null hypothesis of independence, that is, the true copula is the product copula, so that $C = \Pi^d$. Since C_2^n and C_3^n are unbiased estimators of $C^{(2)}$ and $C^{(3)}$, respectively, and using the fact that C_m^n and $C^{(m)}$ have constant densities on the boxes of the uniform partition defined below, we can propose a test based on the distances between $C^{(2)}$ and C_2^n, and $C^{(3)}$ and C_3^n. We consider several different distances, including: the supremum distance, the total variation distance, the Hellinger distance, and even the Kullback-Leibler divergence. This proposal works for every dimension d. Moreover, the exact distributions of the statistics used in the test can be very easily approximated by a large number of simulations, because the sample copula is "computer friendly". Hence, we do not need to employ heavy machinery in order to compute the corresponding null distributions.
We will show, via simulations, that our proposals have acceptable powers in dimensions $d = 2$, and good powers in $d = 3$, and $d = 4$. We note, however, that the tests can be easily extended to higher dimensions.
Also we run the presented tests with real data for dimension 3. We will see that all the test have similar results. The data was taken from the financial markets.
Finally, we give a basic idea for a further investigation. This idea is called exhaustive dependence.

4.1 Introduction

Consider a dimension $d \geq 2$ and consider a d-dimensional random vector X. A very relevant problem is to determine if it is possible to decompose the vector in d independent uni-variate random variables. In some cases we could have evidence that a random vector could be independent and then a statistical test of independence is necessary.

Many tests of statistics already exist. The majority of the tests are based on, at least, one of the following concepts: empirical distribution function, ranks, empirical copula, characteristic function, conditional distribution and distance correlation.

The most studied case is , of course, the case of dimension 2. Several tests have been proposed, for example: the test of Spearman, the test of Hoeffding (1948), the test of Blum-Kieffer-Rosenblatt(1961), and more recently the test of Genest-Remillard (2004) and the test of Bagkavos-Patil (2017), see [12], [3], [7] and [1].

The case $d \geq 3$ has been studied less and there are not as many tests as in the case of dimension 2. We can highlight two reasons for the complications in the case $d \geq 3$: first, the statistics based on the empirical distribution (or in the empirical copula) defined in dimension 2 can be extended easily ti higher dimensions, however, they become difficult to evaluate for large sample sizes; second, for the statistics based on ranks, for example the Spearman's test, there is not a simple extension for the case $d \geq 3$, see [2] and [19].

Our goal is to show that the tests that we propose can be evaluated easily for higher dimensions and for not small sample sizes. Moreover, of course, we have to show that the performance of our tests is good.

4.2 Preliminary Results

We start this section with some basic notions. First, we have a Glivenko-Cantelli's Theorem which gives uniform almost sure convergence of C_m^n to $C^{(m)}$, for every $m \geq 2$, that is,

$$\lim_{n\to\infty} \sup_{(u,v)\in\mathbf{I}^2} |C_m^n(u,v) - C^{(m)}(u,v)| = 0 \text{ a.s.} \tag{4.1}$$

On the other hand, from [4], we also have that

$$\lim_{m\to\infty} \sup_{(u,v)\in\mathbf{I}^2} |C^{(m)}(u,v) - C(u,v)| \leq \lim_{m\to\infty} \frac{d}{2m} = 0. \tag{4.2}$$

From equations (4.1) and (4.2) we get a Glivenko-Cantelli's Theorem for the convergence of C_m^n to C.

Let $\mathbb{P}_{C_m^n}$ and $\mathbb{Q}_{C^{(m)}}$ be the probability measures induced by the sample d-copula C_m^n and by the checkerboard copula $C^{(m)}$, respectively, associated with a d-copula C. Recall that the **total variation distance**, see for example [8], between two probabilty measures $\mathbb{P}$ and $\mathbb{Q}$ on the Borel measurable space $(\mathbf{I}^d, \mathcal{B}(\mathbf{I}^d))$ is defined by

$$d_{TV}(\mathbb{P},\mathbb{Q}) = \sup_{A\in\mathcal{B}(\mathbf{I}^d)} |\mathbb{P}(A) - \mathbb{Q}(A)|. \tag{4.3}$$

Recall that $0 \leq d_{TV} = (\mathbb{P},\mathbb{Q}) \leq 1$.
The total variation distance of $\mathbb{P}$ and $\mathbb{Q}$, in the case that $\mathbb{P}$ and $\mathbb{Q}$ have densities $f_{\mathbb{P}}$ and $f_{\mathbb{Q}}$, with respect to the Lebesgue measure λ^d on the measurable space $(\mathbf{I}^d, \mathcal{B}(\mathbf{I}^d))$, which are constants on the uniform partition of order m of $\mathbf{I}^d$, can be written as

$$d_{TV}(\mathbb{P},\mathbb{Q}) = \sup_{A\in\mathcal{B}(\mathbf{I}^d)} |\mathbb{P}(A) - \mathbb{Q}(A)| = \frac{1}{2}\int_{\mathbf{I}^d} |f_{\mathbb{P}} - f_{\mathbb{Q}}| d\lambda^d. \tag{4.4}$$

Using equations (2.39), (2.14), (4.1) together with equation (4.4), it is easy to see that

Theorem 4.2.1. *Let C be a d-copula, let $m \geq 2$ fixed and let n be a multiple of m, let C_m^n be the sample copula of order m built from a modified sample of size n from C and let $C^{(m)}$ be the checkerboard of order m. If $\mathbb{P}_{C_m^n}$ and $\mathbb{Q}_{C^{(m)}}$ are the probability measures on $\mathbf{I}^d$ defined by C_m^n and $C^{(m)}$ respectively, then*

$$\lim_{n\to\infty} d_{TV}\left(\mathbb{P}_{C_m^n}, \mathbb{Q}_{C^{(m)}}\right) = 0 \quad a.s. \tag{4.5}$$

We give the definition of other important metrics: the **Hellinger distance** and the **supremum distance** or **uniform distance**, see [8]:

Definition 4.2.2. *Let $\mathbb{P}$ and $\mathbb{Q}$ be two probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and let $f_{\mathbb{P}}$ and $g_{\mathbb{Q}}$ be the densities of the measures $\mathbb{P}$ and $\mathbb{Q}$, respectively . We define the* **Hellinger distance** *between $\mathbb{P}$ and $\mathbb{Q}$, denoted by $d_H(\mathbb{P},\mathbb{Q})$, as*

$$d_H(\mathbb{P},\mathbb{Q}) = \frac{1}{\sqrt{2}}\left[\int_{\mathbf{I}^d}\left(\sqrt{f_{\mathbb{P}}} - \sqrt{f_{\mathbb{Q}}}\right)^2 d\lambda^d\right]^{1/2}. \tag{4.6}$$

Note that $0 \leq d_H(\mathbb{P},\mathbb{Q}) \leq 1$.
The Hellinger distance is a type L^2 distance between $\mathbb{P}$ and $\mathbb{Q}$.

Definition 4.2.3. *Let $F_{\mathbb{P}}$ and $F_{\mathbb{Q}}$ be the distribution functions associated with the probability measures $\mathbb{P}$ and $\mathbb{Q}$, respectively. We define the* **supremum distance between $\mathbb{P}$ and $\mathbb{Q}$, denoted by $d_{\mathbf{sup}}(F_{\mathbb{P}}, F_{\mathbb{Q}})$, as**

$$d_{\mathbf{sup}}(F_{\mathbb{P}}, F_{\mathbb{Q}}) = d_\infty(F_{\mathbb{P}}, F_{\mathbb{Q}}) = \sup_{\underline{x}\in\mathbf{I}^d} |F_{\mathbb{P}}(\underline{x}) - F_{\mathbb{Q}}(\underline{x})| \leq 1. \tag{4.7}$$

The sumpremum distance is called the **weak distance**, because it is related to weak convergence.

Finally, we also include another function which is not a metric, called **relative entropy**, also known in Statistics as **Kullback-Leibler divergence**

Definition 4.2.4. *Let two probability measures $\mathbb{P}$ and $\mathbb{Q}$ and let f_P and f_Q the densities of $\mathbb{P}$ and $\mathbb{Q}$, respectively. We define the* ***Kullback-Leibler divergence*** *between $\mathbb{P}$ and $\mathbb{Q}$, denoted by $d_I(\mathbb{P},\mathbb{Q})$, as*

$$d_I(\mathbb{P},\mathbb{Q}) = \int_{S(\mathbb{P})} f_{\mathbb{P}} \log\left(\frac{f_{\mathbb{P}}}{f_{\mathbb{Q}}}\right) d\lambda^d, \tag{4.8}$$

where, $S(\mathbb{P})$ is the support of $\mathbb{P}$ on R^d, and we define $0\log(0/q) = 0$ for every $q \in R$ and $p\log(p/0) = \infty$, see [8].

This divergence satisfies that $d_I(\mathbb{P},\mathbb{P}) = 0$ and $d_I(\mathbb{P},\mathbb{Q}) \geq 0$, but the remaining properties of a metric are not satisfied, because even though $d_I(\mathbb{P},\mathbb{Q}) \in [0,\infty]$, so it can take the value ∞, it is not symmetric, and it does not satisfy the triangle inequality. However, it is an important quantity in Statistics, which measures information gain.
We will use the concepts in equations (4.4), (4.6), (4.7) and (4.8), in the next section to define four statistics to test for multivariate independence.

4.3 Main Theorem

In this section we find a characterization of independence in terms of the checkerboard approximations of a copula.

Theorem 4.3.1. *Let C be a d-copula. Then*

$$C = \Pi_d \textit{ if and only if} C(u_1,\ldots,u_d) = C^{(2)}(u_1,\ldots,u_d) = C^{(3)}(u_1,\ldots,u_d) \textit{ for every } (u_1,\ldots,u_d) \in \mathbf{I}^d, \tag{4.9}$$

where $C^{(2)}$ and $C^{(3)}$ are the checkerboards approximation of the d-copula C of order 2 and 3, respectively.

Proof: First, assume that $d = 2$. Let us assume that $C = \Pi_2$, that is, C is the independence copula, we know that

$$C_m^{'} = \{C(u,v) = u \cdot v \mid u,v \in \{0, 1/m, 2/m, \ldots, (m-1)/m, 1\}\},$$

is a 2-subcopula. For this 2-subcopula and the uniform partition of size m given in equation (2.8), and using equation (1.1), we have that, for every $i_1, i_2 \in I_m$,

$$\begin{aligned} V_{C'_m}(\overline{R^m_{i_1,i_2}}) &= \frac{i_1}{m}\frac{i_2}{m} - \frac{i_1-1}{m}\frac{i_2}{m} - \frac{i_1}{m}\frac{i_2-1}{m} + \frac{i_1-1}{m}\frac{i_2-1}{m} \\ &= \left(\frac{i_1}{m} - \frac{i_1-1}{m}\right)\left(\frac{i_2}{m} - \frac{i_2-1}{m}\right) \\ &= \lambda^2(\overline{R^m_{i_1,i_2}}), \end{aligned} \tag{4.10}$$

where λ^2 is the Lebesgue measure on $(\mathbf{R}^2, \mathcal{B}(\mathbf{R}^2))$.

If we use the bilinear interpolation of Lemma 2.3.5 in Nelsen's book, see [18], we have that $C^{(m)}$ the checkerboard approximation of order m of $C = \Pi_2$ has a density given by equation (2.39)

$$c^{(m)}(u,v) = \frac{V_{C'_m}(\overline{R^m_{i_1,i_2}})}{\lambda^2(\overline{R^m_{i_1,i_2}})} \quad \text{for every} \quad (u,v) \in R^m_{i_1,i_2}, \tag{4.11}$$

for every $i_1, i_2 \in I_m$. On the other hand, using equations (4.10) and (4.11) we have that

$$c^{(m)}(u,v) = \frac{V_{C'_m}(\overline{R^m_{i_1,i_2}})}{\lambda^2(\overline{R^m_{i_1,i_2}})} = \frac{\lambda^2(\overline{R^m_{i_1,i_2}})}{\lambda^2(\overline{R^m_{i_1,i_2}})} = 1 \text{ for every } (u,v) \in R^m_{i_1,i_2},$$

for every $i_1, i_2 \in I_m$. Hence, the density of $C^{(m)}$ is the constant 1 on $\mathbf{I}^2$. Therefore, for every integer $m \geq 2$, the checkerboard approximation $C^{(m)}$ satisfies

$$C^{(m)}(u,v) = \int_0^v \int_0^u 1 ds dt = u \cdot v = \Pi_2(u,v) = C(u,v) \text{ for every } (u,v) \in \mathbf{I}^2. \tag{4.12}$$

In particular this holds for $m = 2$ and $m = 3$.

Now, let us assume that for some 2-copula C we have that $C(u,v) = C^{(2)}(u,v) = C^{(3)}(u,v)$ for every $(u,v) \in \mathbf{I}^2$.

Let $m = 2$ and define $\alpha = V_C([0,1/2]^2) = V_C(R^2_{1,1})$, as in the uniform partition of order $m = 2$, given in equation (2.8). Then, by equation (1.1) and using inequality (1.5), if $\alpha = C(1/2,1/2)$, we have

$$0 = W(1/2,1/2) \leq \alpha = C(1/2,1/2) \leq M(1/2,1/2) = \frac{1}{2}. \tag{4.13}$$

Observe that $R^2_{1,1} \cup R^2_{1,2} = [0,1/2] \times [0,1]$ is a disjoint union. Also, observe that by continuity of C, $V_C(R^2_{1,2}) = V_C(\overline{R^2_{1,2}})$; the same applies to $R^2_{2,1}$ and $R^2_{2,2}$. Hence, using equation (1.1),

$$\frac{1}{2} = V_C([0,1/2] \times [0,1]) = V_C(R^2_{1,1}) + V_C(R^2_{1,2}) = \alpha + V_C(R^2_{1,2})$$

and so $V_C(R^2_{1,2}) = (1/2 - \alpha)$. Similar arguments show that $V_C(R^2_{2,1}) = (1/2 - \alpha)$ and that $V_C(R^2_{2,2}) = \alpha$. So, using the bilinear interpolation we obtain

$$C^{(2)}(u,v) = C_\alpha(u,v) = \begin{cases} 4\alpha uv & \text{if} \quad (u,v) \in R^2_{1,1} \\ 2\alpha u + 4(1/2-\alpha)u(v-1/2) & \text{if} \quad (u,v) \in R^2_{1,2} \\ 2\alpha v + 4(1/2-\alpha)(u-1/2)v & \text{if} \quad (u,v) \in R^2_{2,1} \\ \alpha + (1-2\alpha)(u+v-1) + 4\alpha(u-1/2)(v-1/2) & \text{if} \quad (u,v) \in R^2_{2,2}. \end{cases}$$

From equation (4.14) we have that $C^{(2)}$ is a function of a unique parameter, that is, $\alpha = C(1/2,1/2)$, and from the hypothesis we have that $C(u,v) = C^{(2)}(u,v) = C_\alpha(u,v)$, where from equation (4.13), $0 \leq \alpha \leq 1/2$.
Now, we also assume that C satisfies

$$C_\alpha(u,v) = C(u,v) = C^{(2)}(u,v) = C^{(3)}(u,v) = C^{(3)}_\alpha(u,v), \tag{4.14}$$

for every $(u,v) \in \mathbf{I}^2$. In order to construct $C^{(3)}_\alpha(u,v)$ we need to evaluate all the volumes $V_{C_\alpha}(\overline{R^3_{i_1,i_2}})$ for every $i_1, i_2 \in I_3 = \{1,2,3\}$. We first observe that $R^3_{1,1} = [0,1/3]^2 \subset [0,1/2]^2 = R^2_{1,1}$, so using equation (4.14), we obtain

$$V_{C_\alpha}(R^3_{1,1}) = V_{C_\alpha}([0,1/3]^2) = C_\alpha(1/3,1/3) = \frac{4\alpha}{9}. \tag{4.15}$$

In general, by continuity of C, $V_{C_\alpha}(R^3_{i_1,i_2}) = C_\alpha(i/m, j/m) - C_\alpha((i-1)/m, j/m) - C_\alpha(i/m,(j-1)/m) + C_\alpha((i-1)/m,(j-1)/m)$ for every $i_1, i_2 \in I_3$. We also know from equation (2.8), that $\lambda^2(R^3_{i_1,i_2}) = 1/9$ for every $i_1, i_2 \in I_3$. Hence, using equation (4.11), the density of $C^{(3)}_\alpha$ is given by

$$c^{(3)}_\alpha(u,v) = \frac{V_{C_\alpha}(\overline{R^3_{i_1,i_2}})}{\lambda^2(\overline{R^3_{i_1,i_2}})} = 9V_{C_\alpha}(\overline{R^3_{i_1,i_2}}) \tag{4.16}$$

for every $(u,v) \in R^3_{i_1,i_2}$ and for every $i_1, i_2 \in I_3$. Using equations (4.14) and (4.15) we have that

$$C^{(3)}_\alpha(u,v) = 9V_{C_\alpha}(R^3_{1,1})u \cdot v = 9\left(\frac{4\alpha}{9}\right) u \cdot v = 4\alpha u \cdot v, \tag{4.17}$$

for every $(u,v) \in R^3_{1,1} = [0,1/3]^2$. We also have that

$$\begin{aligned} V_{C_\alpha}(R^3_{1,2}) &= C_\alpha(1/3,2/3) - C_\alpha(1/3,1/3) - C_\alpha(0,2/3) + C_\alpha(0,1/3) \\ &= \frac{2\alpha}{3} + 4\left(\frac{1}{2}-\alpha\right)\left(\frac{1}{3}\right)\left(\frac{1}{6}\right) - \frac{4\alpha}{9} \\ &= \frac{1}{9} + \alpha\cdot\frac{12-4-8}{18} \\ &= \frac{1}{9} \end{aligned} \tag{4.18}$$

Now, using equations(4.17), (4.18) and integration we obtain

$$C_\alpha(u,v) = 9V_{C_\alpha}(R^3_{1,1})u\left(\frac{1}{3}\right) + 9V_{C_\alpha}(R^3_{1,2})u\left(v-\frac{1}{3}\right), \tag{4.19}$$

for every $(u,v) \in R^3_{1,2} = [0,1/3] \times (1/3,2/3]$.

Finally, let us take $(u_0,v_0) = (1/4,1/2)$; then using equation (2.8), we have that $(1/4,1/2) \in (R^2_{1,1} \cap R^3_{1,2})$, and from equation (4.14), we have that

$$C_\alpha(u_0,v_0) = C^{(2)}(1/4,1/2) = 4\alpha\left(\frac{1}{4}\right)\left(\frac{1}{2}\right) = \frac{\alpha}{2}. \tag{4.20}$$

And using equations (4.18) and (4.19), we have that

$$C^{(3)}_\alpha(1/4,1/2) = 4\alpha\left(\frac{1}{4}\right)\left(\frac{1}{3}\right) + 9\left(\frac{1}{9}\right)\left(\frac{1}{4}\right)\left(\frac{1}{6}\right) = \frac{\alpha}{3} + \frac{1}{24}. \tag{4.21}$$

Therefore, from hypothesis (4.14) and equations (4.20) and (4.21), we have that

$$C^{(3)}_\alpha(1/4,1/2) = C_\alpha(1/4,1/2) \text{ if and only if } \alpha/2 = \alpha/3 + 1/24 \text{ if and only if } \alpha = 1/4.$$

But, from equation (4.14), this happens if and only if $C_\alpha(u,v) = \Pi_2(u,v) = u \cdot v$.

We now assume that $d = 3$ and that $C = \Pi_3$ is the product 3-copula. Then we know that

$$C^{'}_m = \{C(u,v,w) = u \cdot v \cdot w \mid u,v,w \in \{0, 1/m, 2/m, \ldots, (m-1)/m, 1\}\},$$

is a 3-subcopula, and for this 3-subcopula and the uniform partition of size m given in equation (2.8), and using equation (1.1), we have by continuity of C that

$$\begin{aligned}V_{C'_m}(R^m_{i_1,i_2,i_3}) &= \frac{i_1}{m}\frac{i_2}{m}\frac{i_3}{m} - \frac{i_1-1}{m}\frac{i_2}{m}\frac{i_3}{m} - \frac{i_1}{m}\frac{i_2-1}{m}\frac{i_3}{m} - \frac{i_1}{m}\frac{i_2}{m}\frac{i_3-1}{m} \\ &\quad + \frac{i_1-1}{m}\frac{i_2-1}{m}\frac{i_3}{m} + \frac{i_1-1}{m}\frac{i_2}{m}\frac{i_3-1}{m} + \frac{i_1}{m}\frac{i_2-1}{m}\frac{i_3-1}{m} - \frac{i_1-1}{m}\frac{i_2-1}{m}\frac{i_3-1}{m} \\ &= \left(\frac{i_1}{m} - \frac{i_1-1}{m}\right)\left(\frac{i_2}{m} - \frac{i_2-1}{m}\right)\left(\frac{i_3}{m} - \frac{i_3-1}{m}\right) \\ &= \lambda^3(R^m_{i_1,i_2,i_3}), \qquad (4.22)\end{aligned}$$

where λ^3 is the Lebesgue measure on $(\mathbf{R}^3, \mathcal{B}(\mathbf{R}^3))$.

If we use the trilinear interpolation of Lemma 2.3.5 in Nelsen's book, [18], we have that $C^{(m)}$ the checkerboard approximation of order m of $C = \Pi_3$ has a density given by equation (2.39)

$$c^{(m)}(u,v,w) = \frac{V_{C'_m}(R^m_{i_1,i_2,i_3})}{\lambda^3(R^m_{i_1,i_2,i_3})} \quad \text{for every} \quad (u,v,w) \in R^m_{i_1,i_2,i_3}, \qquad (4.23)$$

for every $i_1, i_2, i_3 \in I_m$. But, using equations (4.22) and (4.23) we have that

$$c^{(m)}(u,v,w) = \frac{V_{C'_m}(R^m_{i_1,i_2,i_3})}{\lambda^3(R^m_{i_1,i_2,i_3})} = \frac{\lambda^3(R^m_{i_1,i_2,i_3})}{\lambda^3(R^m_{i_1,i_2,i_3})} = 1 \text{ for every } (u,v,w)\ R^m_{i_1,i_2,i_3},$$

for every $i_1, i_2, i_3 \in I_m$. Hence, the density of $C^{(m)}$ is the constant 1 on $\mathbf{I}^3$. Therefore, for every integer $m \geq 2$ the checkerboard approximation $C^{(m)}$ satisfies that

$$C^{(m)}(u,v,w) = \int_0^v \int_0^u \int_0^w 1 ds dt dr = u \cdot v \cdot w = \Pi_3(u,v,w) = C(u,v,w) \text{ for every } (u,v,w) \in \mathbf{I}^3. \qquad (4.24)$$

In particular this holds for $m = 2$ and $m = 3$.

We now prove the converse. Let us assume that for some 3-copula C we have that $C(u,v,w) = C^{(2)}(u,v,w) = C^{(3)}(u,v,w)$ for every $(u,v,w) \in \mathbf{I}^3$.

Let $m = 2$, define $\alpha_0 = V_C([0,1/2]^3) = V_C(R^2_{1,1,1})$, as in the uniform partition of order $m = 2$, given in equation (2.8). Then, by equation (1.1), $\alpha_0 = C(1/2,1/2,1/2)$, and using the inequality (1.5), we have

$$0 = W^3(1/2,1/2,1/2) \leq \alpha_0 = C(1/2,1/2,1/2) \leq M^3(1/2,1/2,1/2) = \frac{1}{2}. \qquad (4.25)$$

Define $\alpha_1 = C(1,1/2,1/2), \alpha_2 = C(1/2,1,1/2)$ and $\alpha_3 = C(1/2,1/2,1)$. Let $C_{1,2}(u,v) = C(u,v,1)$, the we know that $C_{1,2}$ is a 2-copula, and by hypothesis we also know that $C_{1,2}(u,v) =$

$C^{(2)}(u,v,1) = C^{(3)}(u,v,1)$ for every $(u,v) \in \mathbf{I}^2$. It is trivial to see that by linearity in the construction of $C^{(2)}$ and $C^{(3)}$, we have that the checkerboards of $C_{1,2}$ of order $m = 2$ and $m = 3$ are given by $C^{(2)}_{1,2}(u,v) = C^{(2)}(u,v,1)$ and $C^{(3)}_{1,2}(u,v) = C^{(3)}(u,v,1)$ for every $(u,v) \in \mathbf{I}^2$. Therefore, we have the transformed hypotheses

$$C_{1,2}(u,v) = C^{(2)}_{1,2}(u,v) = C^{(3)}_{1,2}(u,v) \quad \text{for every} \quad (u,v) \in \mathbf{I}^2. \tag{4.26}$$

So using what we proved for the case $d = 2$ above, we have that

$$\alpha_3 = C(1/2,1/2,1) = C_{1,2}(1/2,1/2) = \Pi_2(1/2,1/2) = \frac{1}{4}. \tag{4.27}$$

Defining $C_{1,3}(u,w) = C(u,1,w)$ and $C_{2,3}(v,w) = C(1,v,w)$ for every $u,v,w \in \mathbf{I}$, and reasoning as above we observe that

$$\alpha_1 = C(1,1/2,1/2) = C_{2,3}(1/2,1/2) = \frac{1}{4} = C_{1,3}(1/2,1/2) = C(1/2,1,1/2) = \alpha_2. \tag{4.28}$$

Now using the fact that any 3-copula is increasing in each coordinate, together with equations (4.27) and (4.28) and inequality (4.25) we have that

$$0 \leq \alpha_0 = C(1/2,1/2,1/2) \leq \min(\alpha_1,\alpha_2,\alpha_3) = \frac{1}{4}. \tag{4.29}$$

In order to find $C^{(2)}(u,v,w)$, we first need to evaluate the C-volumes of all the uniform boxes $R^2_{i,j,k}$ for every $i,j,k \in I_2$, in order to find its density in each box, which is given by the constant $V_C(\overline{R^2_{i,j,k}})/\lambda^3(R^2_{i,j,k}) = 8 \cdot V_C(\overline{R^2_{i,j,k}})$ for every $i,j,k \in I_2 = \{1,2\}$.

We know that $V_C(\overline{R^2_{1,1,1}}) = C(1/2,1/2,/12) = \alpha_0$. By equation (1.1) and using i) in Definition 1.1.1, we have that $V_C(\overline{R^2_{2,1,1}}) = V_C([1/2,1] \times [0,1/2] \times [0,1/2]) = C(1,1/2,1/2) - C(1/2,1/2,1/2) = \alpha_1 - \alpha_0 = 1/4 - \alpha_0$, similarly $V_C(\overline{R^2_{1,2,1}}) = V_C(\overline{R^2_{1,1,2}}) = 1/4 - \alpha_0$. Again, by equation (1.1) and using i) and ii) in Definition 1.1.1, we obtain $V_C(\overline{R^2_{2,2,1}}) = V_C([1/2,/1] \times [1/2,1] \times [0,1/2]) = C(1,1,1/2) - C(1,1/2,1/2) - C(1/2,1,1/2) + C(1/2,1/2,1/2) = 1/2 - \alpha_1 - \alpha_2 + \alpha_0 = 1/2 - 1/4 - 1/4 + \alpha_0 = \alpha_0$, analogously, $V_C(\overline{R^2_{2,1,2}}) = V_C(\overline{R^2_{1,2,2}}) = \alpha_0$.

Finally, using Definition 1.1.1 we have that $V_C(\overline{R^2_{2,2,2}}) = V_C([1/2,/1] \times [1/2,1] \times [1/2],1) = 1 - 1/2 - 1/2 - 1/2 + 1/4 + 1/4 + 1/4 - \alpha_0 = 1/4 - \alpha_0$.

Therefore, integrating the above density we get $C^{(2)}(u,v,w)$ the checkerboard copula of order $m = 2$, for every $(u,v,w) \in \mathbf{I}^3$, which is given by:

$$C^{(2)}(u,v,w) = \begin{cases} 8\alpha_0 u\cdot v\cdot w & \text{if} \quad (u,v,w)\in R^2_{1,1,1} \\ (2-8\alpha_0)u\cdot v\cdot w + (8\alpha_0-1)u\cdot v & \text{if} \quad (u,v,w)\in R^2_{1,1,2} \\ (2-8\alpha_0)u\cdot v\cdot w + (8\alpha_0-1)u\cdot w & \text{if} \quad (u,v,w)\in R^2_{1,2,1} \\ (2-8\alpha_0)u\cdot v\cdot w + (8\alpha_0-1)v\cdot w & \text{if} \quad (u,v,w)\in R^2_{2,1,1} \\ 8\alpha_0 u\cdot v\cdot w + (1-8\alpha_0)u\cdot v & \\ +(1-8\alpha_0)u\cdot w + (8\alpha_0-1)u & \text{if} \quad (u,v,w)\in R^2_{1,2,2} \\ 8\alpha_0 u\cdot v\cdot w + (1-8\alpha_0)u\cdot v & \\ +(1-8\alpha_0)v\cdot w + (8\alpha_0-1)v & \text{if} \quad (u,v,w)\in R^2_{2,1,2} \\ 8\alpha_0 u\cdot v\cdot w + (1-8\alpha_0)u\cdot w & \\ +(1-8\alpha_0)v\cdot w + (8\alpha_0-1)w & \text{if} \quad (u,v,w)\in R^2_{2,2,1} \\ (1/2-2\alpha_0)\{(u-1/2)+(v-1/2)+(w-1/2)\} & \\ 4\alpha_0\{(u-1/2)(v-1/2)+(u-1/2)(w-1/2)\} & \\ 4\alpha_0(v-1/2)(w-1/2)+\alpha_0 & \\ (2-8\alpha_0)(u-1/2)(v-1/2)(w-1/2) & \text{if} \quad (u,v,w)\in R^2_{2,2,2}. \end{cases}$$

Observe that by hypothesis $C^{(2)}(u,v,w) = C(u,v,w)$, and that by equation (4.30) it has a unique parameter α_0.

In order to obtain $C^{(3)}$, we will obtain its density using equation (4.30), that is,

$$c^{(3)}(u,v,w) = \frac{V_C(\overline{R^3_{i,j,k}})}{\lambda^3(R^3_{i,j,k})} = 27V_{C^{(2)}}(\overline{R^3_{i,j,k}}), \tag{4.30}$$

for every $i,j,k\in I_3$ and for every $(u,v,w)\in R^3_{i,j,k}$, as defined in equation (2.8).

To find the density of $C^{(3)}$ on $R^3_{1,1,1} = [0,1/3]^3$ we observe that $R^3_{1,1,1}\subset R^2_{1,1,1}$, so, using (4.30), $V_{C^{(2)}}(R^3_{1,1,1}) = C^{(2)}(1/3,1/3,1/3) = (8/27)\alpha_0$. To obtain the density of $C^{(3)}$ on $R^3_{1,2,1} = [0,1/3]\times(1/3,2/3]\times[0,1/3]\subset R^2_{1,1,1}\cup R^2_{1,2,1}$, we need $V_{C^{(2)}}(\overline{R^3_{1,2,1}}) = C^{(2)}(1/3,2/3,1/3) - C^{(2)}(1/3,1/3,1/3) = (2-8\alpha_0)(2/27)+(8\alpha_0-1)(1/9)-8\alpha_0(1/27) = 1/27$. For the density of $C^{(3)}$ on $R^3_{1,1,2} = [0,1/3]\times[0,1/3]\times(1/3,2/3]\subset R^2_{1,1,1}\cup R^2_{1,1,2}$ we need $V_{C^{(2)}}(\overline{R^3_{1,1,2}}) = C^{(2)}(1/3,1/3,2/3) - C^{(2)}(1/3,1/3,1/3) = (2-8\alpha_0)(2/27)+(8\alpha_0-1)(1/9)-8\alpha_0(1/27) = 1/27$. Finally, for the density of $C^{(3)}$ on $R^3_{1,2,2} = [0,1/3]\times(1/3,2/3]\times(1/3,2/3]\subset R^2_{1,1,1}\cup R^2_{1,2,1}\cup R^2_{1,1,2}\cup R^2_{1,2,2}$ we need $V_{C^{(2)}}(\overline{R^3_{1,2,2}}) = C^{(2)}(1/3,2/3,2/3) - C^{(2)}(1/3,2/3,1/3) - C^{(2)}(1/3,1/3,2/3) + C^{(2)}(1/3,1/3,1/3) = 8\alpha_0(4/27)+(1-8\alpha_0)(4/9)+(8\alpha_0-1)(1/3)-(2-8\alpha_0)(4/27)-(8\alpha_0-1)(2/9)+8\alpha_0(1/27) = 1/27$. Hence, from equation (4.30), we have that $C^{(3)}$ has density 1 on $R^3_{1,1,2}, R^3_{1,2,1}$ and $R^3_{1,2,2}$, and density $8\alpha_0$ on $R^3_{1,1,1}$.

Let $(u_0,v_0,w_0) = (1/4,1/2,2)\in R^2_{1,1,1}\cap R^3_{1,2,2}$ then by hypothesis $C^{(2)}(1/4,1/2,1/2) = C^{(3)}(1/4,1/2,1/2)$.

Integrating the density of $C^{(3)}$ we have that

$$\begin{aligned} C^{(3)}(1/4,1/2,1/2) &= \int_0^{1/3}\int_0^{1/3}\int_0^{1/4} 8\alpha_0 dudvdw + \int_0^{1/3}\int_{1/3}^{1/2}\int_0^{1/4} dudvdw \\ &\quad + \int_{1/2}^{1/3}\int_0^{1/3}\int_0^{1/4} dudvdw + \int_{1/3}^{1/2}\int_{1/3}^{1/2}\int_0^{1/4} dudvdw \\ &= (2/9)\alpha_0 + (1/72) + (1/72) + (1/144) \\ &= (2/9)\alpha_0 + (5/144). \end{aligned} \tag{4.31}$$

Now, using equation (4.30) we know that $C^{(2)}(1/4,1/2,1/2) = \alpha_0/2$. Therefore, we have that

$$\frac{\alpha_0}{2} = \frac{2}{9}\alpha_0 + \frac{5}{144}.$$

Solving for α_0 we have that $\alpha_0 = 1/8$, and using equation (4.30), we have that $C^{(3)}(u,v,w) = C^{(2)}(u,v,w) = \Pi_2(u,v,w)$ for every $(u,v,w) \in \mathbf{I}^3$.
The rest of the proof follows from an easy induction. □

From equations (4.12) and (4.24) in the last proof we have the following result:

Corollary 4.3.2. *Let $C = \Pi^d$ be the product copula, then for every $m \geq 2$ we have that*

$$C^{(m)}(u_1,\ldots,u_d) = \Pi^d(u_1,\ldots,u_d) \quad \textit{for every} \quad (u_1,\ldots,u_d) \in \mathbf{I}^d.$$

4.4 Independence Tests

The total variation distance defined in equations (4.3) and (4.4) provides the largest possible difference between two probability measures, so it is considered a far stronger distance than the "sup" distance. Many statisticians seem to think that "the d_{TV} is generally too strong to be useful", but this is not so in our case as Theorem 4.2.1 shows.

Using the characterization of independence given in Theorem 4.3.1, we first propose a new independence test based on the total variation distance. We know by equation (4.9) that for $d \geq 2$

$$C = \Pi^d \quad \text{if and only if} \quad C = C^{(2)} = C^{(3)}.$$

Let $Q_{C^{(2)}}$ and $Q_{C^{(3)}}$ be the probability measures induced by the checkerboards of order $m = 2$ and $m = 3$, respectively. Assuming (4.9) holds, if we observe that the probability measure associated to Π^d the product copula is simply the Lebesgue product measure λ^d, then we have that

$$d_{TV}(Q_{C^{(2)}}, \lambda^d) = 0 \quad \text{and} \quad d_{TV}(Q_{C^{(3)}}, \lambda^d) = 0. \tag{4.32}$$

Since the total variation distance is quite strong, we may use it to see whether the true copula C equals the product copula Π^d or not. So, we will use the fact that under H_0, that is, $C = \Pi^d$, we have that $Q_{C^{(2)}}$ and $Q_{C^{(3)}}$ are equal to λ^d by equation (4.32), and besides, by Theorem 4.2.1, $Q_{C^{(2)}}$ and $Q_{C^{(3)}}$ are the uniform limits of $P_{C_2^n}$ and $P_{C_3^n}$ as n increases. Hence, based on Corollary 4.3.2 we propose the statistic

$$\eta_{TV}(C; n) = \frac{d_{TV}(P_{C_2^n}, \lambda^d) + d_{TV}(P_{C_3^n}, \lambda^d)}{2}. \tag{4.33}$$

In this case, we take a sample from a true copula C with sample size n, and $P_{C_m^n}$ for $m = 2$ and $m = 3$ are the probability measures induced by the sample copulas C_m^n with orders $m = 2$ and $m = 3$, respectively. Since in our case, the alternative hypothesis is $H_1 : C \neq \Pi^d$, then we have that under H_0, by equation (4.5), $\lim_{n\to\infty} \eta_{TV}(C; n) = 0$ almost surely. Also for any copula $C \neq \Pi^d$ we have that $\lim_{n\to\infty} \eta_{TV}(C; n) > 0$.

Even if the null distribution of the test statistic, $\eta_{TV}(\Pi^d; n)$ is not known for a fixed sample size n, it is straightforward to generate a large number of simulations in a very reasonable time, even for not so small dimension d and sample sizes n, in order to approximate the quantiles, let us say of order $90\%, 95\%$ and 99%, needed in order to perform a standard test. Of course, we reject H_0 at levels $\alpha = 0.10, 0.05$ and $\alpha = 0.01$, if the observed value of $\eta_{TV}(C; n)$ exceeds the respective $(1 - \alpha)$ quantiles above.

We can also use different distances, other than the total variation distance. For example, we can use the Hellinger distance given in equation (4.6), or the supremum distance given in equation (4.7). Furthermore, we can even use the Kullback-Leibler divergence in equation (4.8). Since the densities of the product copulas Π^d, and the sample d-copulas of order m are constant on the d-boxes of the uniform partition as in equation (2.8), d_I satisfies that $d_I(Q_{C^{(2)}}, \lambda^d) = 0 = d_I(Q_{C^{(3)}}, \lambda^d)$ if $C = \Pi^d$; however both are greater than zero if $C \neq \Pi^d$. Therefore, we can also use any of the following statistics to test for multivariate independence.

$$\eta_H(C; n) = \frac{d_H(P_{C_2^n}, \lambda^d) + d_H(P_{C_3^n}, \lambda^d)}{2}, \tag{4.34}$$

$$\eta_{sup}(C; n) = \frac{d\text{sup}(C_2^n, \Pi^d) + d\text{sup}(C_3^n, \Pi^d)}{2}. \tag{4.35}$$

$$\eta_I(C; n) = \frac{d_I(P_{C_2^n}, \lambda^d) + d_I(P_{C_3^n}, \lambda^d)}{2}. \tag{4.36}$$

When the sample is not a multiple of six, we must modify the sample size. We need some criterion to, we remove some data. Some criteria may be: remove the most recent data, remove the oldest data, remove the data randomly, etc. Let n be the size of the original sample, then we define the sample size after removing the data, denoted by n^* as

$$n^* = 6\left\lfloor \frac{n}{6} \right\rfloor.$$

It is relevant to note that at most five data may be removed. When the sample size is big, removing five data does not impact the results significantly

4.5 Simulations

In this Section we will carry out a simulation study in dimensions $d = 2$, $d = 3$ and $d = 4$. We start by doing a comparison to several well known proposals of tests of independence, in the case $d = 2$. For dimensions $d = 2$ and $d = 3$ we present the results of the comparison among our statistics mentioned above for several different families of copulas. We also discuss some tests of independence in the multivariate case.

For example, we simulate samples for a Clayton copula with $\rho = 0.05$ and for the product copula. We run the simulations 1,000 times for sample sizes equal to 60, 600 and 6000, see Figure 4.1 and Figure 4.3 ; then we present the results of some basic statistics such as the mean, the maximum and the minimum the grids generated by the uniform partitions of size 2 and 3, see Figure 4.2 and Figure 4.4. The graphs of both simulations are very similar for every sample size, since ρ is very close to 0. Instead of that, the statistics presented differ. The data corresponding to the mean must be interpreted in the same way as the partition of $\mathbf{I}^2$, generated by the uniform partitions of size 2 and 3. Of course, the mean in each rectangle, in the case of the product copula, is very close to 1/4 for $m = 2$, and is close to 1/9 for $m = 3$. For the Clayton copula, the values of the mean are a little far from the values 1/4 and 1/9, for $m = 2$ and $m = 3$, respectively. Now, for the minimum and the maximum, the values for the independent copula are closer than the values of the Clayton copula to 1/4 and 1/9.

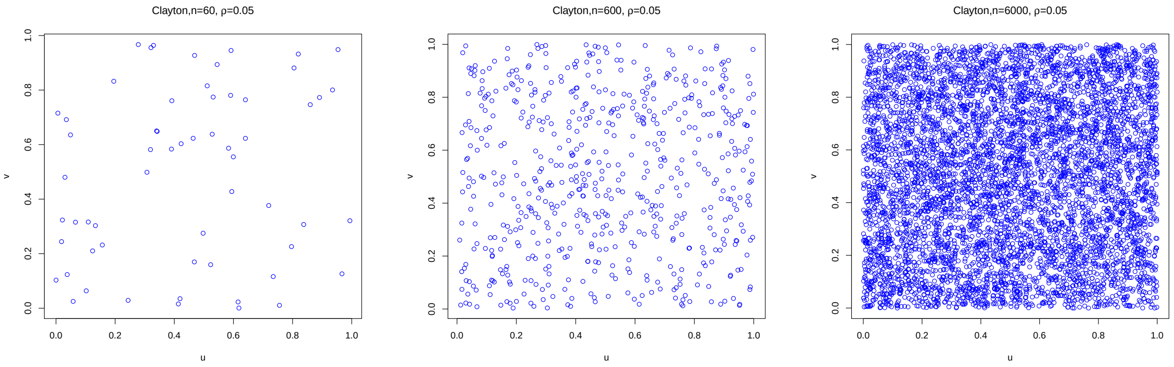
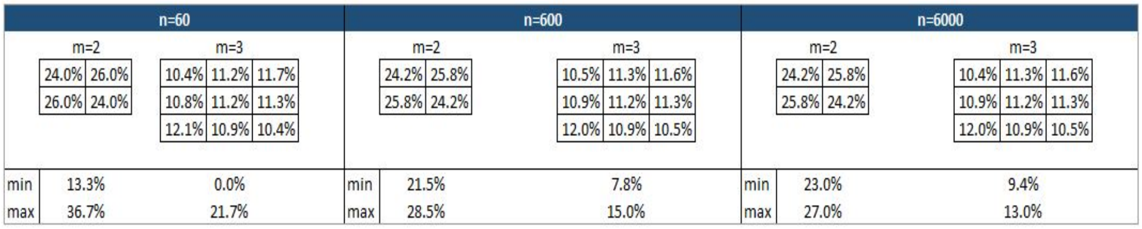


Figure 4.1: Clayton copula. Sample size of 60, 600 and 6000, respectively, and $d = 2$



n=60						n=600						n=6000					
	m=2		m=3				m=2		m=3				m=2		m=3		
	24.0%	26.0%	10.4%	11.2%	11.7%		24.2%	25.8%	10.5%	11.3%	11.6%		24.2%	25.8%	10.4%	11.3%	11.6%
	26.0%	24.0%	10.8%	11.2%	11.3%		25.8%	24.2%	10.9%	11.2%	11.3%		25.8%	24.2%	10.9%	11.2%	11.3%
			12.1%	10.9%	10.4%				12.0%	10.9%	10.5%				12.0%	10.9%	10.5%
min	13.3%		0.0%			min	21.5%		7.8%			min	23.0%		9.4%		
max	36.7%		21.7%			max	28.5%		15.0%			max	27.0%		13.0%		

Figure 4.2: Statistics for a Clayton copula. Sample size of 60, 600 and 6000, respectively, and $d = 2$

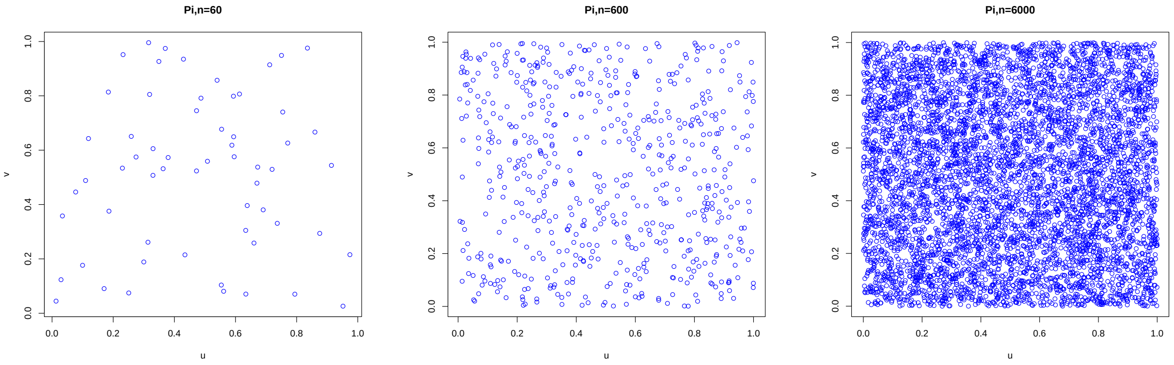
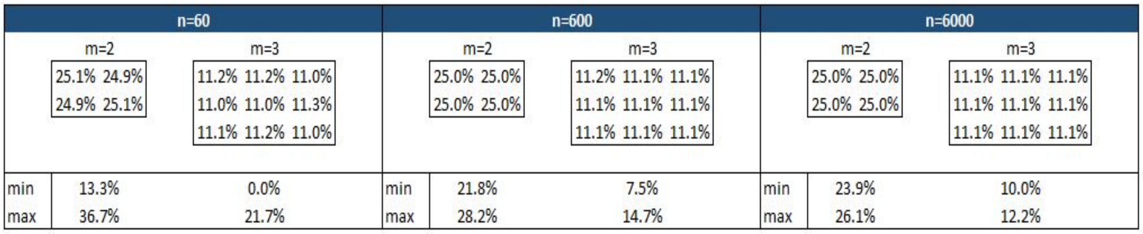


Figure 4.3: Product copula. Sample size of 60, 600 and 6000, respectively, and $d = 2$



	n=60					n=600					n=6000				
	m=2		m=3			m=2		m=3			m=2		m=3		
	25.1%	24.9%	11.2%	11.2%	11.0%	25.0%	25.0%	11.2%	11.1%	11.1%	25.0%	25.0%	11.1%	11.1%	11.1%
	24.9%	25.1%	11.0%	11.0%	11.3%	25.0%	25.0%	11.1%	11.1%	11.1%	25.0%	25.0%	11.1%	11.1%	11.1%
			11.1%	11.2%	11.0%			11.1%	11.1%	11.1%			11.1%	11.1%	11.1%
min	13.3%		0.0%			21.8%		7.5%			23.9%		10.0%		
max	36.7%		21.7%			28.2%		14.7%			26.1%		12.2%		

Figure 4.4: Statistics for the product copula. Sample size of 60, 600 and 6000, respectively, and $d = 2$

4.5.1 Dimension $d = 2$

In this subsection we will study the case $d = 2$, which has been the most studied case for independent tests. Several statistics have been proposed to test for independence. For example, we have two classical tests; the first one proposed by Hoeffding in 1948, to test the independence of two continuous random variables with continuous joint and marginal densities, see [12]. This test is based on the function $D(x, y) = F(x, y) - F(x, \infty) \cdot F(\infty, y) = F(x, y) - F_1(x) \cdot F_2(y)$, where F denotes the joint distribution function, F_1 and F_2 are the margins of X and Y, and $\Delta(F) = \int D^2(x, y) dF(x, y)$. The statistic he proposed is based on the joint empirical distribution function minus the product of the marginal empirical functions. Here we used the **hoeffd** function of the R PACKAGE HMISC. The second test is based on extensions of this result and is known as the Blum-Kiefer-Rosenblatt's independence tests, see [3]. See also [17] for null Gaussian approximations of the bivariate Blum-Kiefer-Rosenblatt (BKR) test of independence. To perform the BKR test we use their statistic B_n and the normal approximation in [17]. Since the results of these two statistics are always quite similar, here we only report the results based on Hoeffding's statistic.

Another well known test for independence in the case $d = 2$ is based on Spearman's ρ, and has been used extensively in applications. Here we use the **spearman.test** function of the R PACKAGE PSPEARMAN. However, it is well known that this test has low power if the distribution of the alternative is continuous but singular, as is the case for several copulas.

We used a small value of the sample size, $n = 36$, to compare our results to other works which also use small sample sizes for their simulations.

We first observe that, if we are simulating from a standard Archimedean copula, such as Clayton, Gumbel, Frank, etc., the power obtained by using the tests of Hoeffding, Blum-Kiefer-Rosenblatt and Spearman's ρ are a little better at levels $\alpha = 0.01, 0.05$ and $\alpha = 0.10$ than the ones we obtain using the statistics given in equations (4.33), (4.34) and (4.36); that is, the total variation, the Hellinger distances and the Kullback-Leibler divergence, see Figure 4.5. It is important to note that most of all these Archimedean copulas are absolutely continuous, with complete support and with smooth densities.

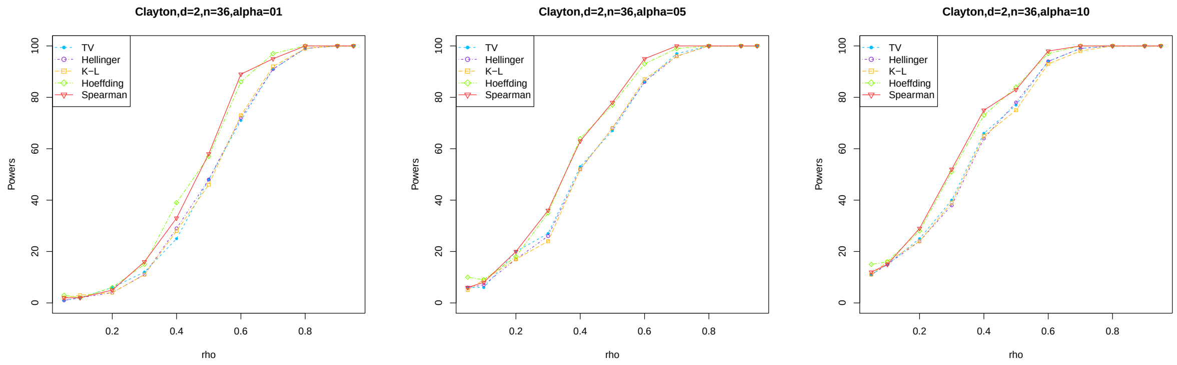


Figure 4.5: Powers for the Clayton family with $n = 36$ in dimension $d = 2$

It is not difficult to see, via simulations, that the independent tests of Hoeffding and Blum-Kiefer-Rosenblatt have a problem with small sample sizes. In fact, we note that if we are sampling from the independent copula Π_2, and test at the usual levels $\alpha = 0.01, 0.05$ and $\alpha = 0.10$, the real levels of the test do not correspond to the desired values of α. For example, if we set $\alpha = 0.05$ and perform several simulations, the actual value of α under independence is approximately $\alpha = 0.075$. Something similar happens with the other two values of α. This happens because there is an effect of discretization of the statistic when the sample size is small. Therefore, we recommend caution when using these two tests with small sample sizes. For a better detail, observe Figure 4.6.

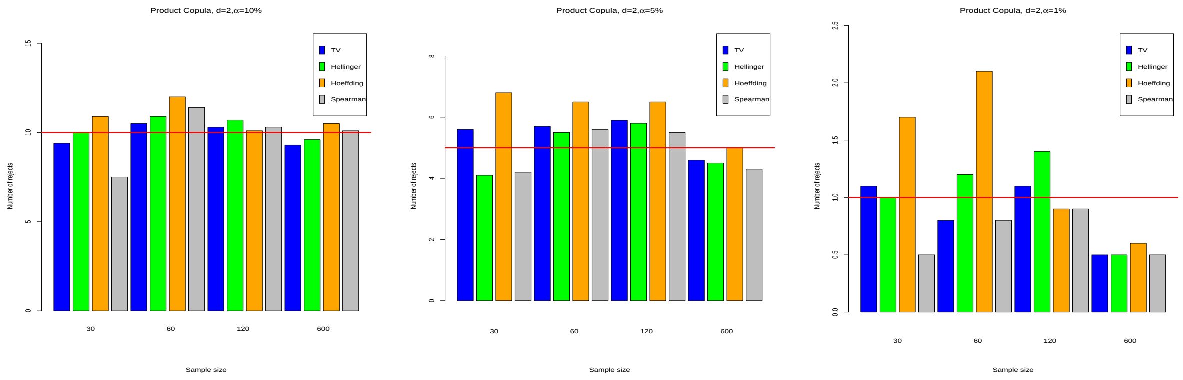


Figure 4.6: Powers for the Product copula with dimension $d = 2$

We did not use the supremum distance in the simulations because we observed a strong discretization effect of the statistic (4.35); that is, the different values observed from this statistic were very limited, with many ties.

As a second example, we use the Fréchet-Mardia copulas. In this case, we use a convex mixture of W_2 and M_2, the Fréchet-Hoeffding bounds, for the Figure 4.7, and we use a convex mixture of W_2, M_2 and Π_2 for the Figure 4.8. Remember that the copulas defined in Figure 4.7 are singular and the copulas defined in Figure 4.8 are absolutely continuous. As we can see in Figure 4.7 and Figure 4.8, the Spearman's test has very low power, specially for singular copulas. We also note that the total variation statistic in equation (4.33) performs a little better than the Hoeffding and the Blum-Kiefer-Rosenblatt tests at the three levels, but the statistics given in equations (4.34) and (4.36) have the best performance at all three levels, and have a power really close to 100% when $\alpha = 0.05$ and $\alpha = 0.10$.

In the Figure 4.8, the parameters in the convex combination $aM_2 + bW_2 + (1 - a - b)\Pi_2$ begin with $a = 0.5$ and $b = 0.5$, and after b reduces until 0 while $a = 0.5$; after that, a reduces until 0 and b remains equal to 0.

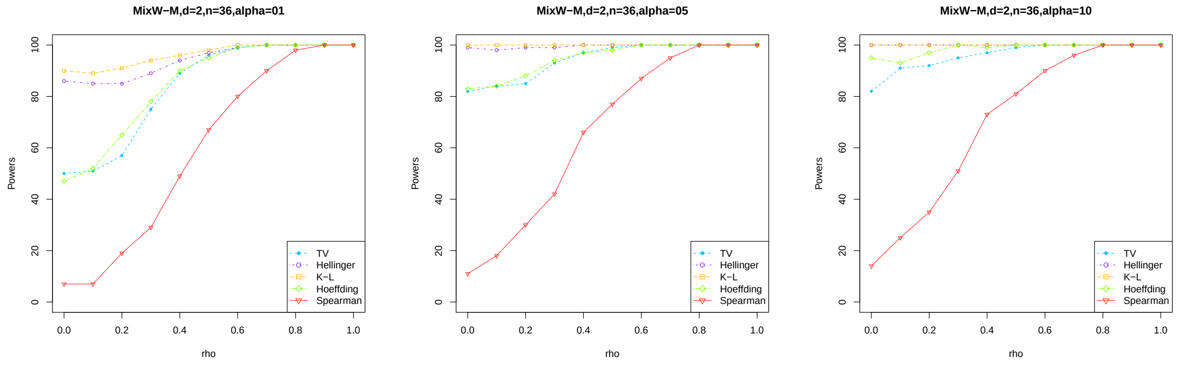


Figure 4.7: Powers for the Frechet-Mardia family with $n = 36$ in dimension $d = 2$ and $\Pi = 0$

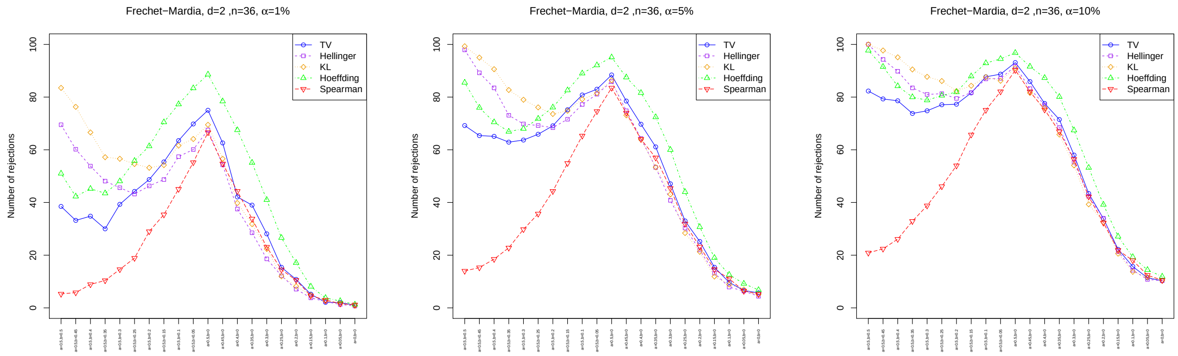


Figure 4.8: Powers for the Frechet-Mardia family with $n = 36$ in dimension $d = 2$ and $\Pi \neq 0$

Finally, we use a convex combination of a Gumbel and a Gumbel-ID, where the latter denotes a

Gumbel distribution with an increasing transformation in its first coordinate and a decreasing transformation in its second coordinate. That is, we used the transformation $(U,V) \to (U, 1-V)$, for any observation (U_i, V_i) of the Gumbel copula. The notation ID stands for *increasing-decreasing* transformation. As we can see in Figure 4.9, in this case the Spearman's ρ, the Hoeffding and the Blum-Kiefer-Rosenblatt have lower powers than our three statistics (4.33), (4.34) and (4.36). Note, however, that in this case the powers for these three statistics may be far better than the ones obtained using the standard tests. In particular, the Kulback-Leibler test (4.36) has the highest powers.

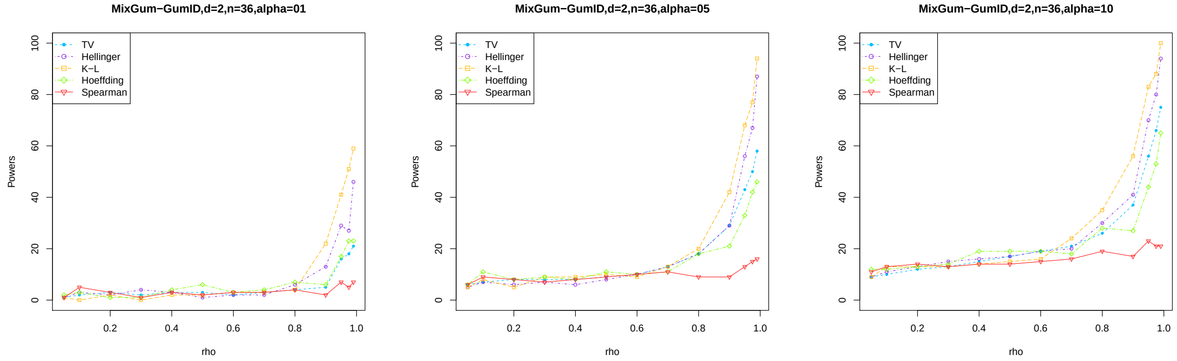


Figure 4.9: Powers for the Mixture Gumb.-Gumb.ID family with $n = 36$ in dimension $d = 2$

Additionally, we include three examples from three absolutely continuous copulas: Clayton, Gumbel and normal, see Figure 4.10, Figure 4.11 and Figure 4.12. The sample size in this case is 60. We can observe that the results are very similar as in the case of sample size equal to 36.

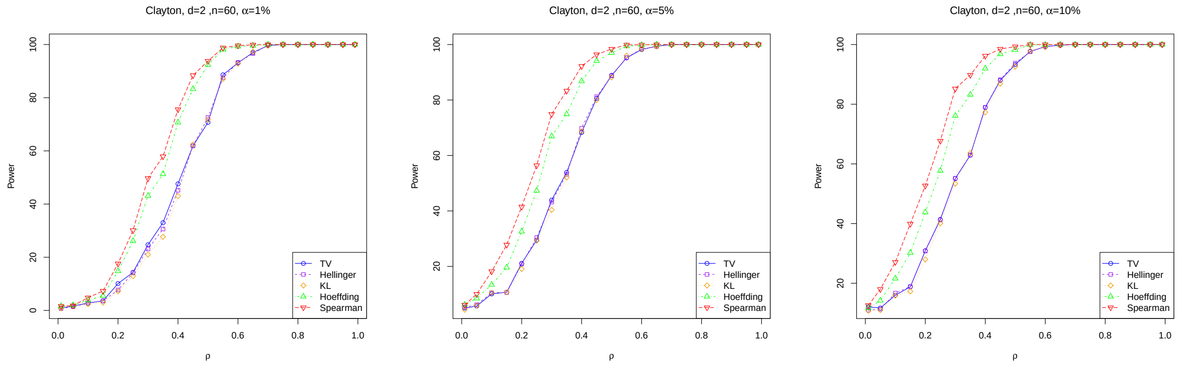


Figure 4.10: Powers for the Clayton family with $n = 60$ in dimension $d = 2$

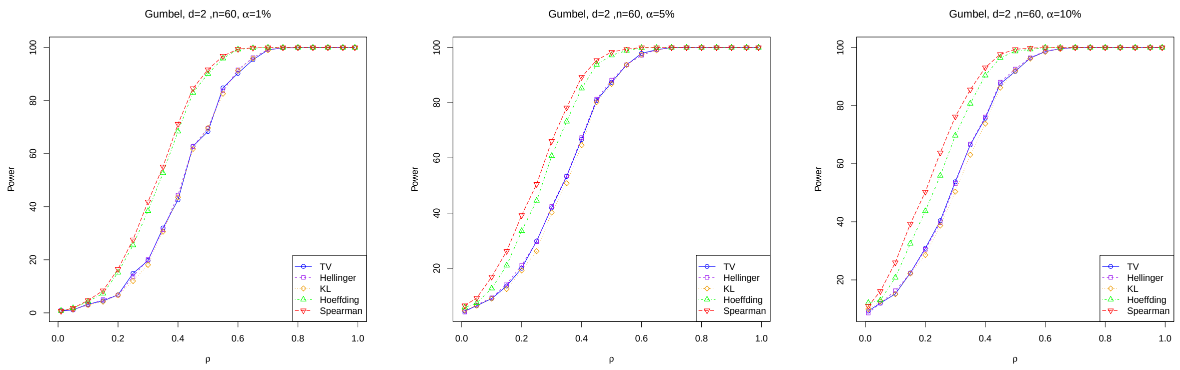


Figure 4.11: Powers for the Gumbel family with $n = 60$ in dimension $d = 2$

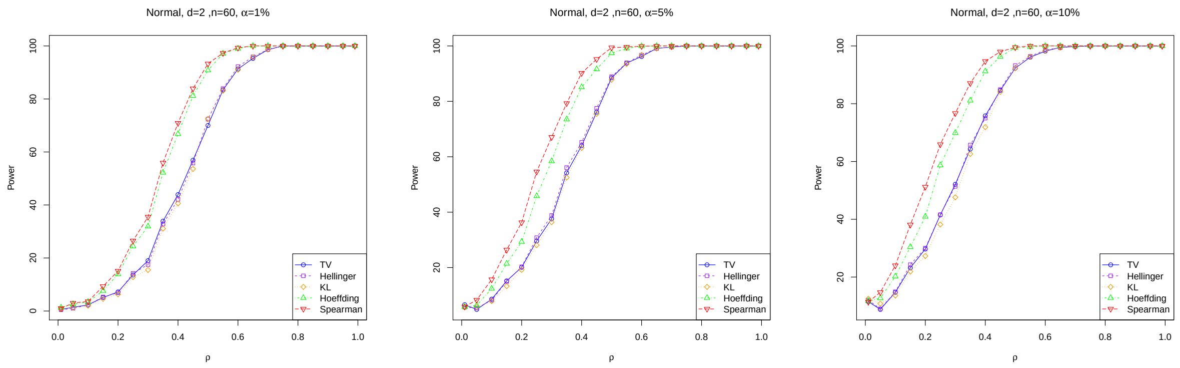


Figure 4.12: Powers for the Normal family with $n = 60$ in dimension $d = 2$

Hence, we observe that our statistics (4.33), (4.34) and (4.36) are competitive even in the case $d = 2$. Besides, in the case of a very smooth copula with complete support and $0.5 < |\hat{\rho}| < 0.95$, where $\hat{\rho}$ is the estimated value of Spearman's rho, we found that for small values of the sample size n, the Spearman's ρ test had the best powers.

In a recent paper, see [1], the authors proposed a new test of bivariate independence based on the idea that if (X, Y) is a random vector with joint distribution F, then under independence $F_Y(y|x) = F_Y(y)$ and $F_Y^{-1}(y|x) = F_Y^{-1}(y)$, that is, they do not depend on the value of X. Here $F_Y(y|x)$ denotes the c.d.f of Y conditional on $X = x$ and $F_Y^{-1}(y|x)$ is its inverse. Based on this idea they construct a new statistic T_n based on the empirical joint distribution function. They compare their statistic to six different tests, including them the Hoeffding D_n, the BKR B_n and the Spearman's rank statistic S_n. In Figure 1 and Figure 2 of [1], with sample size $n = 40$, they see that their T_n is competitive for some alternatives, and does not present a power to far below in the case of copulas when the sample size $n = 60$. We did not include this statistic in our study, but from the figures in [1], we can see that our proposals are close to the values of S_n, and are always competitive if the joint distribution function is continuous and smooth.

4.5.2 Dimension $d = 3$

Many of the statistics proposed in dimension $d = 2$ for independence tests may be extended to higher dimensions. In particular, some extensions to dimension $d = 3$ are somehow natural. For example, in the case of the Hoeffding statistic, we could obviously use $D(x, y, z) = F(x, y, z) - F_1(x) \cdot F_2(y) \cdot F_3(z)$, where F denotes the joint distribution function and F_1, F_2 and F_3 are the margins of X, Y and Z, respectively. If we define $\Delta(F) = \int D^2(x, y, z)dF(x, y, z)$ we could use the empirical version D_n of the previous Δ in order to test for independence. The Blum-Kiefer-Rosenblatt statistic could be similarly extended to dimension $d = 3$. Many other statistics based on the empirical copula have 3-dimensional versions, for example the statistic G_n of Genest and Rémillard, [7]. The problem with all these possible extensions is that in dimension $d = 3$ the empirical copula becomes unfeasible if the sample size is not small. For example if $n = 1000$ then the array necessary to obtain the empirical copula is of size 10^9, which blocks a lot of the memory in a standard computer. Also since we have to operate with it to evaluate several of these statistics, it becomes useless to try in large simulation studies. Note that, even in the case of dimension $d = 2$, most of the papers that have been written proposing new independence tests deal only with simulations based on small sample sizes, in most cases $n \leq 100$. The reason is that, even in dimension $d = 2$, the evaluation of some of the statistics become prohibitively slow if the sample size is moderately large. On the other hand, every statistician who has used the empirical distribution function for data in dimensions greater than or equal to $d = 3$, knows that they need large sample sizes in order to obtain reasonable approximations of the true distribution function F using the empirical distribution function F_n. In the case of 3-copulas, the same is true with the empirical (sub)copula C_n.

In some cases, one can find the asymptotic distribution of a statistic based on the empirical distribution function or empirical copula, but the limiting distribution can only be reached with large extremely sample sizes. In such cases, the statistic is unfortunately impossible to evaluate using a standard computer. Hence, it is not possible to assess for which values of the sample size n the limiting distribution is actually reached.

There are another tests, based on the empirical process or multivariate characteristic functions, which also have problems when working with large sample sizes. See for example [6].

In our simulations using our proposals given in equations (4.33), (4.34), (4.35) and (4.36), we use large values of the sample size n. In many instances the other tests take a very long time, or are even impossible to evaluate. Therefore, we only have compared our proposals among themselves, in order to see which one of them has better power in each of the different cases.

In [2], the authors propose a statistic, $\mathcal{I}_n^2$, which coincides with the square product moment correlation when $d = 2$. The power of this test is good only for absolutely continuous random variables, and it has the same problem as the Spearman's ρ test in dimension $d = 2$.

In [19], the authors give a new test of multivariate independence based on analogues to Kendall's tau and Spearman's ρ. The comments given in the previous paragraph also apply to these tests.

In Figures 4.13 through 4.15 we analyze the case $d = 3$. We used three sample sizes $n = 60$, $n = 120$ and $n = 216$, with $N = 10000$ simulations to find the critical values of the tests under H_0. We also generated a thousand simulations under the alternative H_1, to find the powers of the four tests based in the statistics given in equations (4.33), (4.35), (4.34) and (4.36).

In Figure 4.13 we consider the Gumbel family, with $\alpha = 0.10$. We observe that for n small the statistic based on the supremum distance has better powers, but for $n = 216$ the powers are similar for all tests. It is important to note that the statistic based on the supremum distance has a strong discretization effect, which means that the critical values for this test are not very accurate. The same effect is observed in Figure 4.14.

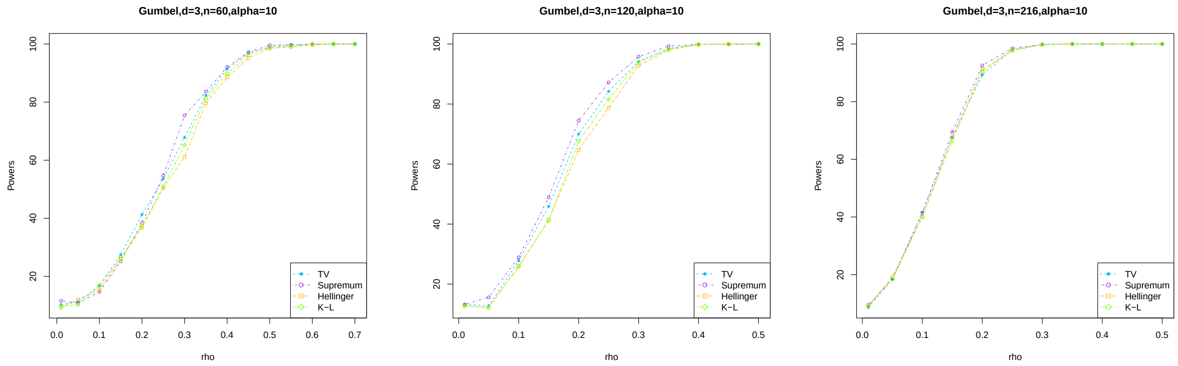


Figure 4.13: Powers for the Gumbel family with different n in dimension $d = 3$

In Figure 4.14 we consider the normal family where the covariance matrix has equal correlations for each pair of variables. We also observe that the supremum distance has a little better power, that disipates when the sample size increases.

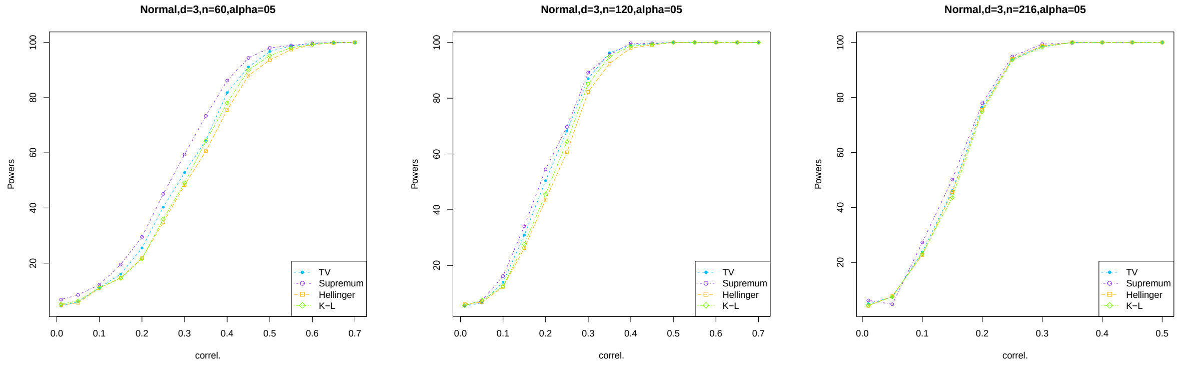


Figure 4.14: Powers for the Normal family with different n and R in dimension $d = 3$

In Figure 4.15 we also study the normal family, but now one of the variables is independent of the other two. In this case we observe that the supremum distance has the worse power compared with the other three statistics, but for large values of n this difference dilutes.

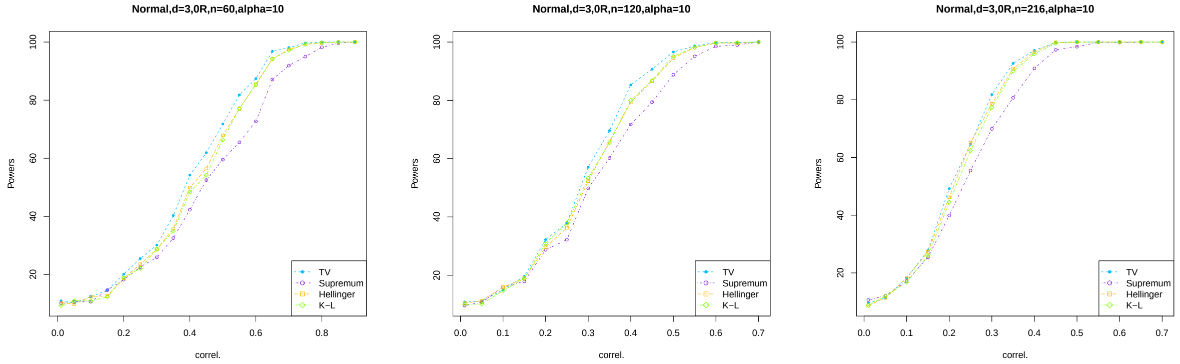


Figure 4.15: Powers for the Normal family with different n and R weak independence in dimension $d = 3$

4.5.3 Dimension $d = 4$

The comments made for the case $d = 3$ also apply to the case $d = 4$. The only difference is that we now take different values of the sample size which include $n = 600$ and $n = 1296$. (The value $1296 = (16) \cdot (81)$ is obtained by multiplying the number of boxes of $C^{(2)}$ and $C^{(3)}$ in dimension $d = 4$). It is important to observe that if we try to evaluate the empirical distribution function of a sample in dimension $d = 4$ and sample size $n = 1296$, in a computer we would only get an error message because the array needed to get it is of size $(1296)^4 = 2821109907456$, which no personal computer can manage.

In Figure 4.16 we consider the Frank family. With $\alpha = 0.05$, we have the same remarks as in Figure 4.13.

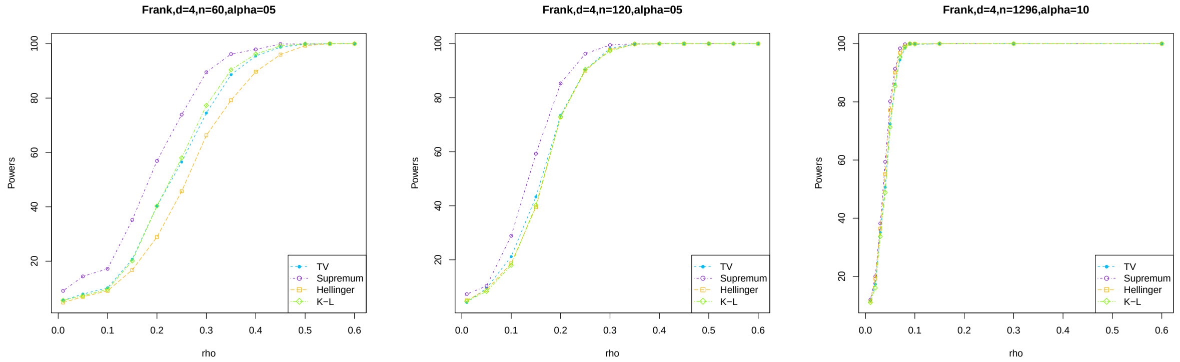


Figure 4.16: Powers for the Frank family with different n in dimension $d = 4$

In Figure 4.17 we study the normal distribution with same correlation among all the random variables, whereas in Figure 4.18 one random variable is independent of the other three, giving weak dependence, The results are quite similar to those given in Figures 4.14 and 4.15.

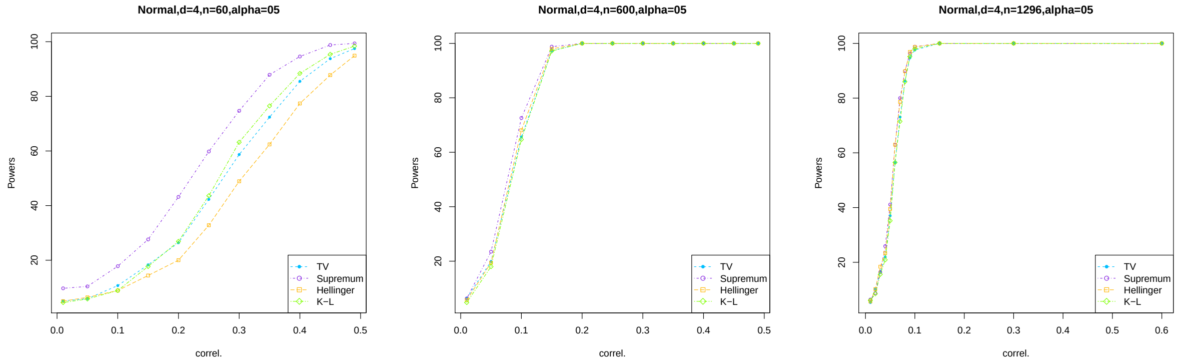


Figure 4.17: Powers for the Normal family with different n and R in dimension $d = 4$

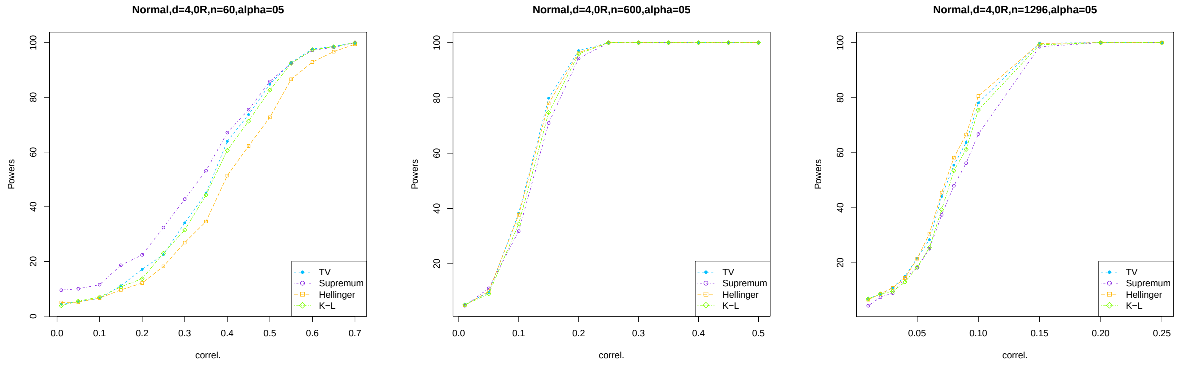


Figure 4.18: Powers for the Normal family with different n and R weak indep. in dimension $d = 4$

Finally, we study the t distribution with same correlation among all the random variables. As we can see in Figure 4.19, for a sample size of $n = 60$, the statistic based on the supremum distance has the highest power followed by the statistic based in the Kullback-Liebler divergence, but this behavior changes for larger sample sizes, for example, in the case of a sample size $n = 1296$, the statistic based on the supremum distance has the worst performance and the statistic based in the divergence has the best.

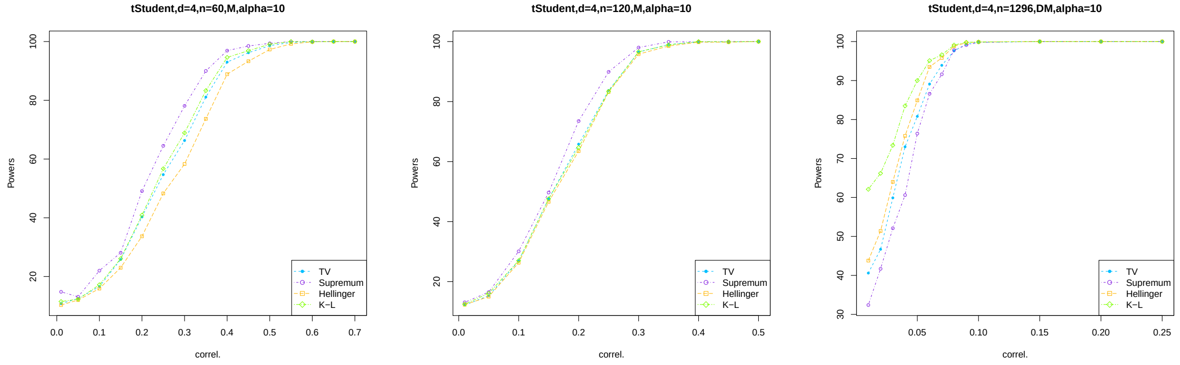


Figure 4.19: Powers for the t distribution with different n and same correlation in dimension $d = 4$

4.6 Real data

In this section we present the results of the proposed tests using real data. In the financial world, a problem of high relevance is to determine if a set of financial variables are independent. For example, we can ask if there is dependence between the reference rate and a certain index of the equity market, or between a certain foreign exchange rate and the prices of the oil, etc.

For this illustration, we use three important variables in the Mexican financial market: the index of the Bolsa Mexicana de Valores, called IPC (Index of Price and Quotations), the exchange rate for the US Dollar to the Mexican Peso (USDMXN), and the price of the sovereign bond with term of one year called MBono. The data that we used correspond to the closing price taken daily from 02/01/2015 to 13/02/2018 (dd/mm/yyyy). The data was taken from the Bloomberg platform with the variable called *px_last*. It is important to highlight that in the period that we are considering there were many moments of high volatility, and the most relevant were: the Brexit referendum in June 2016, the presidential election in the USA in November 2016 and the sharp fall of the equity market presented in February 2018. The sample size observed was $n = 780$.

As we can see in the graphics, the IPC index and the MXNUSD currency present an upward drift; on the other side, the value of the MBono presents a downward drift, see 4.20.

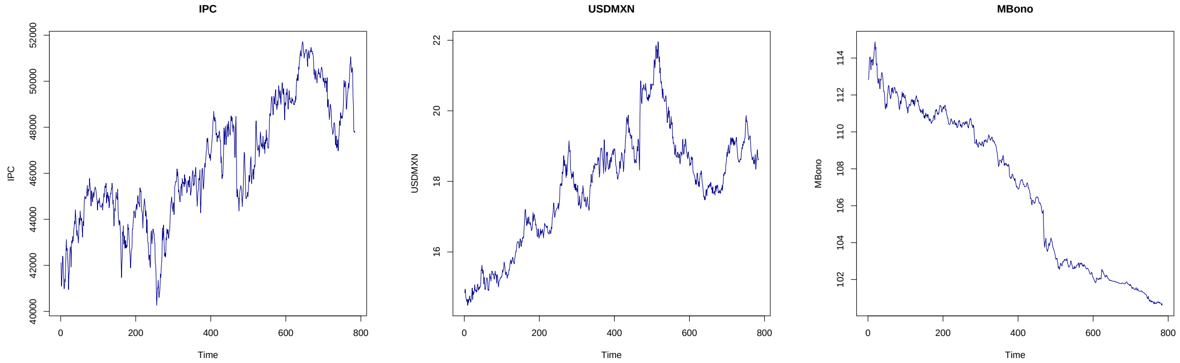


Figure 4.20: Time series of daily values of the IPC, USDMXN and MBono

As is usual in finance, we work with the returns (arithmetic) instead of the original values of the variables in order to center and to stabilize them.

The variance of the returns for the IPC index is bounded and it seems that it is not time dependent. For the case of the USDMXN exchange rate, the variance is stationary and the majority of the data is close to the media, with a few exceptions. The variance of the MBono is obviously time dependent, as we can see in the third graph of Figure 4.21; it is high at the beginning and it reduces practically to a constant at the end.

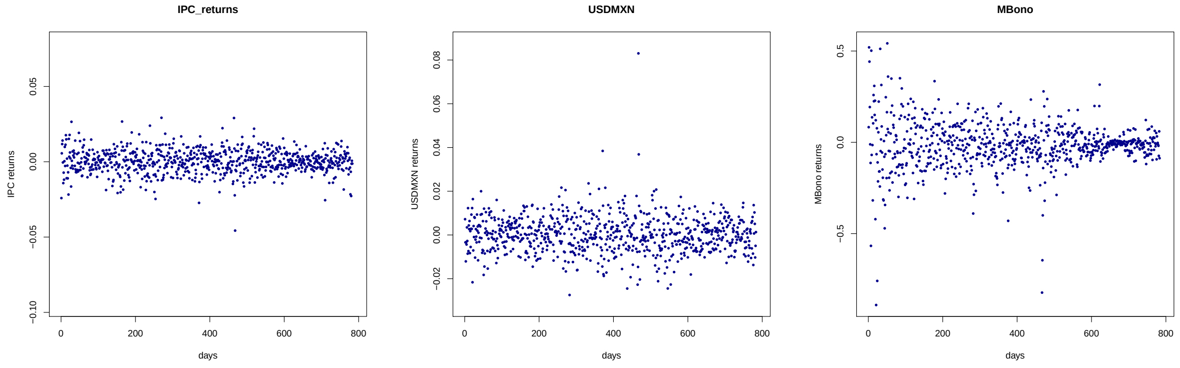


Figure 4.21: Time series of daily returns of the IPC, USDMXN and MBono

We run the tests using all possible pairs and the joint distribution of the 3 variables, that is, we run the tests in the bivariate case for the IPC vs USDMXN, IPC vs MBono and USDMXN vs MBono, and for the case of 3 dimensions, IPC vs USDMXN vs MBono.

In Figure 4.22 we show the modified sample for the bivariate cases. In the first graph, we can notice that there is dependence between the IPC index and the USDMXN exchange rate because there is a higher concentration of the modified sample in the corners of the second diagonal of the unit box. In the second graph, the related one to the IPC index and the MBono, we can not clearly appreciate a concentration of the modified sample in a particular place of the unit box. For the third graph, we have a similar case to the first graph, that is, we can appreciate higher concentration in two corners.

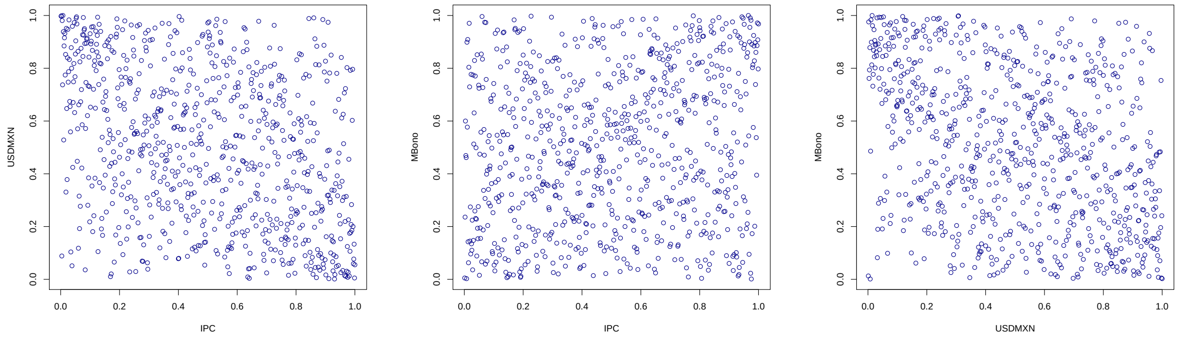


Figure 4.22: Modified sample for the returns IPC vs USDMXN, IPC vs MBono and USDMXN vs MBono, respectively

According to the tests, the variables IPC, USDMXN and MBono are not independent two to two for the significance levels of 0.01, 0.05 and 0.1. In the following tables we can observe the

p-value for each test. We can notice that for every case, the p-value is 0 for the distance of total variation, for the Hellinger distance and for the Kulback- Liebler divergence; it is equal to 0 in two cases for the Spearman; and it is close to 0 for Hoeffding and B-K-R statistics.

IPC vs USDMXN	
Test	p**-value**
Total variation	0
Hellinger	0
Kulback-Liebler	0
Hoeffding	0.00000001
BKR	0.00000271
Spearman	0

IPC vs MBono	
Test	p**-value**
Total variation	0
Hellinger	0
Kulback-Liebler	0
Hoeffding	0.00000166
BKR	0.00024956
Spearman	0.00001147

USDMXN vs MBono	
Test	p**-value**
Total variation	0
Hellinger	0
Kulback-Liebler	0
Hoeffding	0.00000001
BKR	0.00000298
Spearman	0

The following graph presents the modified sample of the returns of the IPC index, USDMXN currency and MBono. We can see that the sample is not uniformly distributed in $\mathbf{I}^3$, see Figure 4.23.

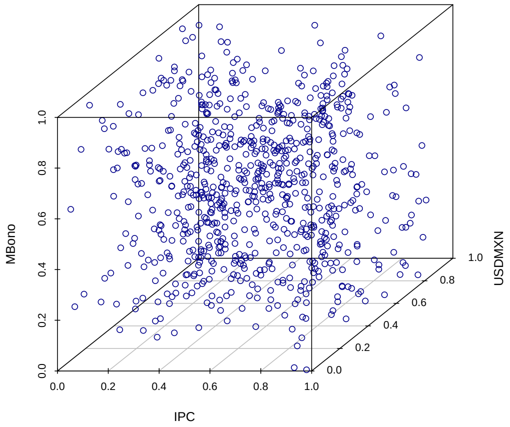


Figure 4.23: Modified sample of the returns for IPC vs USDMXN vs MBono

Now, for the tridimensional case, we reject independence in all the tests for the levels 0.01, 0.05 and 0.1. Moreover, the p-value is 0 for all the statistics.

4.7 Applications

A further investigation based on the work presented in this thesis is the following: What is the maximum separation in which we can divide a vector in independent subvectors? To answer this question we need the following definition

Definition 4.7.1. *Let $\underline{X} = (X_1, \ldots, X_d)$ be a d-dimensional random vector. We say that $\underline{X}$ is* **exhaustively dependent** *if and only if the distribution function of $\underline{X}$ can not be decomposed in the product of the distribution functions associated to any independent subvectors of $\underline{X}$.*

For example, we consider a random vector $\underline{X} = (X_1, X_2, X_3, X_4)$. Then $\underline{X}$ is exhaustively dependent if its distribution function can not be expressed in any of the following forms:

$$F_{X_1,X_2,X_3,X_4}(x_1, x_2, x_3, x_4) = F_{X_{\sigma(1)}}(x_{\sigma(1)})F_{X_{\sigma(2)},X_{\sigma(3)},X_{\sigma(4)}}(x_{\sigma(2)}, x_{\sigma(3)}, x_{\sigma(4)}),$$

$$F_{X_1,X_2,X_3,X_4}(x_1, x_2, x_3, x_4) = F_{X_{\sigma(1)},X_{\sigma(2)}}(x_{\sigma(1)}, x_{\sigma(2)})F_{X_{\sigma(3)},X_{\sigma(4)}}(x_{\sigma(3)}, x_{\sigma(4)}),$$

$$F_{X_1,X_2,X_3,X_4}(x_1, x_2, x_3, x_4) = F_{X_{\sigma(1)}}(x_{\sigma(1)})F_{X_{\sigma(2)}}(x_{\sigma(2)})F_{X_{\sigma(3)},X_{\sigma(4)}}(x_{\sigma(3)}, x_{\sigma(4)}),$$

$$F_{X_1,X_2,X_3,X_4}(x_1, x_2, x_3, x_4) = F_{X_{\sigma(1)}}(x_{\sigma(1)})F_{X_{\sigma(2)}}(x_{\sigma(2)})F_{X_{\sigma(3)}}(x_{\sigma(3)})F_{X_{\sigma(4)}}(x_{\sigma(4)}),$$

where σ is any permutation of $I_4 = \{1, 2, 3, 4\}$, $F_{X_1,X_2,X_3,X_4}(x_1, x_2, x_3, x_4)$ is the joint distribution function of $\underline{X}$ and $F_{X_{\sigma(1)},\ldots X_{\sigma(k)}}$ for $k \leq 3$ are the marginals.

Then, the tests proposed in this thesis can be used to determine if a vector is not exhaustively dependent and hence can be decomposed in independent subvectors.

Exhaustive dependence may be very useful in many applications. For example, in finance it is relevant to identify if there is dependence between a group of variables. However, in many real examples it is known from experience that some of the variables of interest are dependent. Hence we can test if the assumption holds or not. For example, if we consider a vector $\underline{X} = (X_1, X_2, X_3, X_4, X_5)$ that is not exhaustively dependent, and suppose that it can be expressed as

$$F_{\underline{X}}(x_1, x_2, x_3, x_4, x_5) = F_{X_1,X_2}(x_1, x_2)F_{X_3,X_4,X_5}(x_3, x_4, x_5),$$

if the assumption is not rejected then it may be easier to model the structure of dependence between (X_1, X_2) and between (X_3, X_4, X_5) than the structure of dependence of $\underline{X}$.

Chapter 5

Final Comments

In the second chapter of this thesis we presented the concept of sample d-copula and we showed that it has important properties such as it has constant density and a version of a Glivenko-Cantelli's theorem. Additionally, we compare the sample d-copula against the empirical copula, and we noticed that there are many advantages using the sample d-copula instead the empirical. In the third chapter we showed that it is possible to obtain the exact distribution of the sample frequencies under the assumption of independence. We found simple expressions for the distribution of the frequencies, and also we found expressions for the means and covariances. At the end of the chapter we noticed that, in the case of dimension 2, the distribution in the cases of the partition of order 2 and 3 is determined for only one and four random variables, respectively. The main result of this thesis is the trivial characterization of multivariate independence in terms of the checkerboards of order $m = 2$ and $m = 3$, $C^{(2)}$ and $C^{(3)}$. As we have seen, C_m^n the sample d-copula of order m give us a very easy way to estimate the checkerboards of order $m = 2$ and $m = 3$, C_m^n is an estimator of the true copula C, and also, an estimator of $C^{(m)}$, for every $2 \leq m \leq n$. Besides, this estimator can be evaluated with a standard computer, even for large sample sizes and in dimension not so small, which allows us to perform a large number of simulations in order to estimate the distributions under independence of the statistics that we propose; this in contrast to some of the most famous tests that are impossible to evaluate for large sample sizes in high dimensions.

Our proposed statistics are defined in terms of various distances and the Kulback-Leibler divergence. Therefore, we do need not at all the heavy machinery of asymptotic theory to perform the tests.

We simulated examples in dimension $d = 2$, $d = 3$ and $d = 4$, with different models and different sample sizes, including a sample size of $n = 1296$ in dimension $d = 4$. As pointed out above, it would be impossible to compute the empirical distribution function for these values of n and d. Hence, any statistic based on this function is useless in this case, even if we rely on asymptotic results. In [10], there are interesting results about the rate of the weak convergence of the sample process.

As we could observe the tests that we proposed are very competitive in comparison with the most used tests in dimension 2; in fact, for small sample sizes we recommend to use the Hoeffding's test and the Blum's test with caution because it has problems with the true value of α. Besides, the Spearman's test present problems with singular continuous random vectors.

In our simulations we observed that if the sample size is moderately large, any of the statistics that we proposed may be used because they have similar powers. However, when the sample size is small we warn the user against the statistic based on the supremum distance, because it is affected by a strong discretization, which causes problems with the real values of α, the probability of type I error. For more detail of this, observe the following graphs for dimension $d = 3$, clearly the power of the test based on the supremum distance are higher than the correspondent levels, see Figure 5.1:

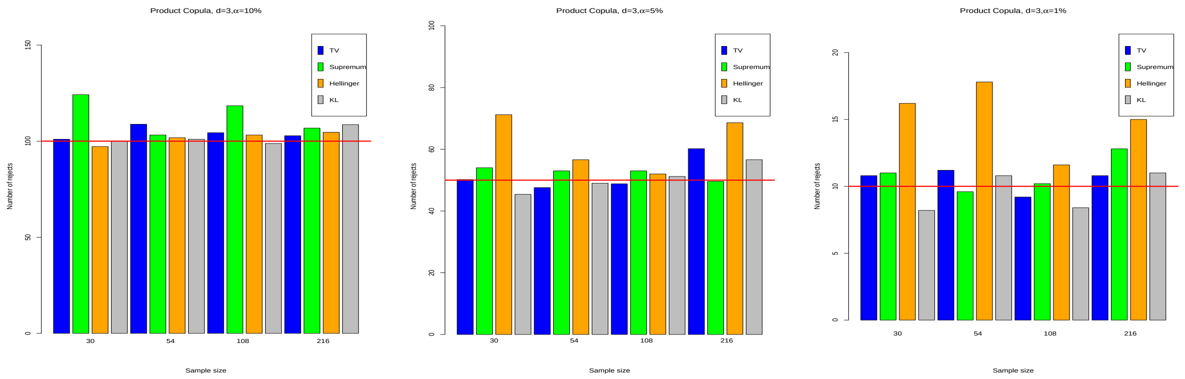


Figure 5.1: Powers for the Product copula with dimension $d = 3$

Bibliography

[1] Bagkavos, D. and Patil, P.N. (2017). A new test of independence for bivariate observations. *J. Multivariate Anal.*, **160**, 117-133.

[2] Bakirov, N.K., Rizzo, M.L. and Székely, G.J. (2006). A multivariate nonparametric test of independence. *J. Multivariate Anal.*, **97**, 1742–1756.

[3] Blum, J.R., Kiefer, J. and Rosenblatt, M. (1961). Distribution free tests of independence based on the sample distribution function. *Ann. Math. Statist.*, **32**, 485–498.

[4] Cuberos, A., Masiello, E. and Maume-Deschamps, V. (2016). Copulas checker-type approximations: applications to quantiles estimation of aggregated variables. ⟨hal-012201838v2⟩.

[5] Durante, F. and Fernández-Sánchez, J. (2010). Multivariate shuffles and approximations of copulas. *Statist. Probab. Lett.*, **80**, 1827–1834.

[6] Fan, Y., Lafaye de Micheaux, P., Penev, S. and Salopek, D. (2017). Multivariate nonparametric test of independence. *J. Multivariate Anal.*, **153**, 189–210.0

[7] Genest, C. and Rémillard, B. (2004). Tests of independence and randomness based on the empirical copula process. *TEST*, **13**, 335-369.

[8] Gibbs, A.L. and Su, F.E. (2002). On choosing and bounding probability metrics. arXiv:math/020902v1[math.PR]3Sep2002, 1–21.

[9] González-Barrios, J.M. and Hernández-Cedillo, M.M. (2013). Sample d-copula of order m. *Kybernetika*, **49**, 663–669.

[10] González-Barrios, J.M. and Hoyos-Argüelles, R. (2016). Distributions associated to the counting techniques of the d-sample copula of order m and weak convergence of the sample process. *Preimpreso*, **167**, IIMAS, UNAM.

[11] González-Barrios, J.M. and Hoyos-Argüelles, R. (2017). Sample copula, Bernstein copula and empirical copula. *Preimpreso*, **169**, IIMAS, UNAM.

[12] Hoeffding, W. (1948). A nonparametric test of independence. *Ann. Math. Statist.*, **19**, 546–557.

[13] Jansen, P., Swanepoel, J. and Veraverbeke, N. (2012). Large sample behavior of the Bernstein copula estimator. *J. Statist. Plann. Inference*, **142**, 1189-1197.

[14] Li, X., Mikusiński, P. and Taylor, M.D. (1998). Strong approximations of copulas. *J. Math. Anal. Appl.*, **225**, 608–623.

[15] Li, X., Mikusiński, P., Sherwood, H., and Taylor, M.D. (1997), *On approximation of copulas.* In Benes, V. and Stepán, J., editors, *Distributions with given marginals and moment problems.*

[16] Mikusiński, P. and Taylor, M.D. (2010) Some approximations of n-copulas. *Metrika*, **72**, 385–414.

[17] Mudholkar, G.S. and Wilding, G.E. (2005). Two Wilson-Hilfetry type approximations for the null distribution of the Blum, Kiefer and Rosenblatt test of bivariate independence. *J. Statist. Plann. Inference*, **128**, 31–41.

[18] Nelsen, R.B. (2006). *An introduction to copulas.* Second ed., Lect. Notes in Statist., Springer-Verlag, New York.

[19] Taskinen, S., Randles, R.H. and Oja, H. (2005). Multivariate nonparametric tests of independence. *J. Amer. Statist. Assoc.*, **100**, (471), 916–925.

[20] Segers, J., Sibuya, M. and Tsukahara, H. (2017). The empirical beta copula. *J. Multivariate Anal.*, **155**, 35–51.