ANNALES
POLONICI MATHEMATICI
96.1 (2009)

Markov operators and n-copulas

by P. Mikusiński and M. D. Taylor (Orlando, FL)

Abstract. We extend the definition of Markov operator in the sense of J. R. Brown and of earlier work of the authors to a setting appropriate to the study of n-copulas. Basic properties of this extension are studied.

1. Introduction. In [3], J. R. Brown introduced Markov operators as positive operators $T : L^\infty(X) \to L^\infty(X)$ satisfying $T1 = 1$ and $\int_X Tf\, d\nu = \int_X f\, d\nu$ where $(X, \mathcal{A}, \nu)$ was a given probability space. One of his main interests was the role played by the particular operators $T_\phi f = f \circ \phi$ induced by invertible measure-preserving $\phi : X \to X$. Another was in the fact that under the constraints he imposed, there was a one-to-one correspondence between the set of Markov operators T on X and the set of doubly stochastic measures μ on $X \times X$. (To say that μ was doubly stochastic in this setting meant that for measurable sets A of X, one had $\mu(A \times X) = \mu(X \times A) = \nu(A)$.)

It would appear that an inspiration for Brown's definition of Markov operator was the theory of Markov processes. One might think of $T : L^\infty(X) \to L^\infty(X)$ as describing the "evolution", $f \mapsto Tf$, of a function over a fixed time interval. However, we found Brown's work interesting because of its applicability to a different mathematical scenario.

Brown required his probability measure ν to be nonatomic, which implied, by results on Borel equivalences (see, for example, [14]), that one could take X to be the unit interval, $I = [0, 1]$, ν could be taken to be λ, 1-dimensional Lebesgue measure on I, and μ would be a doubly stochastic measure on I^2. A *2-copula* is a function $C : I^2 \to I$ which is related to a doubly stochastic measure μ on I^2 by $C(x_1, x_2) = \mu([0, x_1] \times [0, x_2])$. ($C$ may be regarded as the joint distribution function of two random variables that

2000 *Mathematics Subject Classification*: Primary 47N30; Secondary 60E05, 62H05.
Key words and phrases: copula, Markov operator, conditional expectation.

DOI: 10.4064/ap96-1-7

are uniformly distributed over I.) Then one has the following:

$$\text{2-copulas} \overset{\text{1-1}}{\longleftrightarrow} \begin{Bmatrix}\text{doubly stochastic} \\ \text{measures on } I^2\end{Bmatrix} \overset{\text{1-1}}{\longleftrightarrow} \begin{Bmatrix}\text{Markov operators} \\ \text{on } L^\infty(I)\end{Bmatrix}.$$

In [8]–[10], advantage is taken of the correspondence between 2-copulas and Markov operators to introduce and study a type of sequential convergence of 2-copulas. There is also a noteworthy connection between Markov operators and a certain product of 2-copulas that is introduced in [5] and that is of interest in the study of Markov processes: If A and B are 2-copulas, the 2-copula $A * B$ is their product, and T_A, T_B, and T_{A*B} are the associated Markov operators, then

$$T_{A*B} = T_B T_A.$$

Most work on copulas has been with 2-copulas. However, n-copulas for $n \geq 2$ are also of interest. They can be used in modelling systems involving random variables and in studying dependence relations for n-tuples of random variables. (See [13] and [15] for an overview.) Some instances of a desire to push the use of copulas into higher dimensions are [2], [6], [12], and [16] where copulas are used to investigate measures of concordance.

It is the goal of this paper to generalize the concept of a Markov operator in a way that will be useful for the study of n-copulas. An important difference from the previously considered 2-dimensional case is that since we have many different factorizations $I^n = I^p \times I^q$ of the n-dimensional cube, we have many different Markov operators $T : L^\infty(I^p) \to L^\infty(I^q)$ associated with a given n-copula C. Some particular topics that we examine are these:

1. The association of Markov operators with integral kernels and partial derivatives of copulas.
2. "Joining" or "composing" copulas to obtain new copulas.
3. Convergence of copulas. The papers [7]–[10] use Markov operators to study this topic for 2-copulas. For this topic we draw on [11] which investigates approximations of n-copulas.

2. Markov operators. First, a few words about the general setting:

If $(X, \mathcal{X})$, and $(Y, \mathcal{Y})$ are measurable spaces, then $X \times Y$ is automatically considered to be the measurable space with the product σ-algebra $\mathcal{X} \times \mathcal{Y}$. We shall rarely bother to mention $\mathcal{X}$ or $\mathcal{Y}$ and shall just assume their existence.

We shall usually denote probability measures on X and Y as μ_X and μ_Y respectively. If μ is a probability measure on $X \times Y$, we say that μ_X and μ_Y are the *marginals* of μ provided

$$\mu_X(R) = \mu(R \times Y) \quad \text{and} \quad \mu_Y(S) = \mu(X \times S)$$

whenever $R \in \mathcal{X}$ and $S \in \mathcal{Y}$.

Assume that μ_X and μ_Y are the marginals of μ. If $f : X \to \mathbb{R}$ is a measurable function, then, in a trivial way, we may consider that $f : X \times Y \to \mathbb{R}$ by identifying f with f' such that $f'(x, y) = f(x)$ where $x \in X$ and $y \in Y$. Further, we may consider that $L^p(X)$ "$\subseteq$" $L^p(X \times Y)$ for $1 \le p \le \infty$. In the case where $1 \le p < \infty$, this follows from the fact that for $f \in L^p(X)$ it can easily be seen that $\int_X f\, d\mu_X = \int_{X\times Y} f\, d\mu$. Indeed, for such f we have equality of L^p-norms, $\|f\|_X = \|f\|_{X\times Y}$, so we are justified in simply writing $\|f\|$. Similarly we have $L^p(Y)$ "$\subseteq$" $L^p(X \times Y)$.

Now we are ready for Markov operators.

Definition 1. A linear operator $T : L^\infty(X) \to L^\infty(Y)$ is called a *Markov operator* if

(a) $T1 = 1$,
(b) for every $f \in L^\infty(X)$, $f \ge 0$ implies $Tf \ge 0$,
(c) $\int_Y Tf\, d\mu_Y = \int_X f\, d\mu_X$ for every $f \in L^\infty(X)$.

Note that (a) and (b) imply that Markov operators are bounded. Another basic property of Markov operators is that we can just as well take them as being defined on all the L^p-spaces for $1 \le p \le \infty$.

Theorem 1. *Let $T : L^\infty(X) \to L^\infty(Y)$ be a Markov operator. For every $p \ge 1$, T has an extension to a bounded operator $T : L^p(X) \to L^p(Y)$.*

Proof. Let $f \in L^\infty(X)$. Since

$$\int_Y |Tf|\, d\mu_Y \le \int_Y T|f|\, d\mu_Y = \int_X |f|\, d\mu_X$$

and $L^\infty(X)$ is dense in $L^1(X)$, T has a unique extension to a bounded operator $T : L^1(X) \to L^1(Y)$. Hence, by the Riesz–Thorin theorem, T has a unique extension to a bounded operator $T : L^p(X) \to L^p(Y)$ for every $1 < p < \infty$. ∎

Example 1. Here is perhaps the simplest example of a Markov operator: Let $\phi : Y \to X$ be a measurable map where the probability measure μ_Y on Y is given and the probability measure μ_X on X is defined by

$$\mu_X(S) = \mu_Y(\phi^{-1}(S)).$$

Then the map $T : L^\infty(X) \to L^\infty(Y)$ defined by $Tf = f \circ \phi$ is easily verified to be a Markov operator.

Example 2. A *doubly stochastic measure* is a probability measure ν on the unit square, $I^2 = [0, 1]^2$, having the property that $\nu(A\times I) = \nu(I \times A) =$ the one-dimensional Lebesgue measure of A whenever A is a Borel subset of $I = [0, 1]$. In this example we construct a Markov operator that induces an analogue to a checkerboard approximation of a doubly stochastic measure.

(See, for example, [4] and [8] for checkerboard approximations of doubly stochastic measures. We also later define the checkerboard approximation to an n-copula.)

Consider the probability spaces (X, μ_X) and (Y, μ_Y). Let $R_1, \ldots, R_k \subseteq X$ and $S_1, \ldots, S_m \subseteq Y$ be partitions of X and Y into subsets of positive measure. Define $r_0 = s_0 = 0$ and

$$r_i = \sum_{n=1}^{i} \mu_X(R_n) \quad \text{and} \quad s_j = \sum_{n=1}^{j} \mu_Y(S_n)$$

for $i = 1, \ldots, k$ and $j = 1, \ldots, m$. Then $0 = r_0 < r_1 < \cdots < r_k = 1$ and $0 = s_0 < s_1 < \cdots < s_m = 1$. Let ν be a doubly stochastic measure on I^2. For $i = 1, \ldots, k$ and $j = 1, \ldots, m$ define

$$A_{ij} = \frac{\nu([r_{i-1}, r_i] \times [s_{j-1}, s_j])}{\mu_X(R_i)\mu_Y(S_j)}.$$

It is straightforward to verify that the operator $T : L^\infty(X) \to L^\infty(Y)$ defined by

$$Tf = \sum_{i=1}^{k} \sum_{j=1}^{m} A_{ij} \Big(\int_{R_i} f \, d\mu_X \Big) \chi_{S_j}$$

is a Markov operator.

Note that the restriction to $R_{i_0} \times S_{j_0}$ of the measure μ induced by T (see Theorem 2 below) on $X \times Y$ is a product measure, for any $i_0 = 1, \ldots, k$ and $j_0 = 1, \ldots, m$. Indeed, if $A \subseteq R_{i_0}$ and $B \subseteq S_{j_0}$ are Borel sets, then

$$\mu(A \times B) = A_{i_0 j_0} \mu_X(A) \mu_Y(B).$$

This is analogous to the way checkerboard approximations of doubly stochastic measures are constructed.

The next theorem exhibits the most important property of a Markov operator, its relation to a probability measure on $X \times Y$.

THEOREM 2. *Assume μ_X and μ_Y are probability measures on X and Y respectively. Let $T : L^\infty(X) \to L^\infty(Y)$ be a Markov operator. For measurable sets $R \subseteq X$ and $S \subseteq Y$ define*

$$\mu(R \times S) = \int_Y (T\chi_R)\chi_S \, d\mu_Y. \tag{1}$$

Then μ can be extended to a probability measure on $X \times Y$ having μ_X and μ_Y as its marginals.

If, on the other hand, μ is a probability measure on $X \times Y$ having μ_X and μ_Y as marginals, then there is a unique Markov operator $T : L^\infty(X) \to$

$L^\infty(Y)$ such that

$$\int_Y (Tf)g\,d\mu_Y = \int_{X\times Y} fg\,d\mu \tag{2}$$

for all $f \in L^\infty(X)$ and $g \in L^\infty(Y)$. Furthermore, this T is related to μ by (1).

Proof. By Theorem 1.2.8 of [1], μ can be extended to a measure if we can show that it is "continuous" on an increasing sequence of rectangles. Let $R_1 \times S_1 \subseteq R_2 \times S_2 \subseteq \cdots \subseteq X \times Y$ be a sequence of measurable sets and let $R \times S = \bigcup_{j=1}^\infty R_j \times S_j$. The definition of a Markov operator implies $T\chi_{R_n} \to T\chi_R$ μ_Y-a.e. Hence $(T\chi_{R_n})\,\chi_{S_n} \to (T\chi_R)\,\chi_S$ μ_Y-a.e. and, since the sequence is increasing,

$$\mu(R_n \times S_n) = \int_Y (T\chi_{R_n})\chi_{S_n}\,d\mu_Y \to \int_Y (T\chi_R)\chi_S\,d\mu_Y = \mu(R\times S).$$

Thus μ extends to a (σ-additive) measure.

It is easily seen that the marginals of μ are μ_X and μ_Y.

Now suppose we are given a probability measure μ on $X \times Y$ with marginals μ_X and μ_Y.

Let $f \in L^2(X)$ and $g \in L^2(Y)$. By the Cauchy–Schwarz inequality, we have

$$\Big|\int_{X\times Y} fg\,d\mu\Big| \le \Big(\int_{X\times Y} f^2\,d\mu\Big)^{1/2}\Big(\int_{X\times Y} g^2\,d\mu\Big)^{1/2} = \|f\|_X\|g\|_Y. \tag{3}$$

Thus $l_f : L^2(Y) \to \mathbb{R}$ given by $l_f(g) = \int_{X\times Y} fg\,d\mu$ defines a bounded linear functional. By the Riesz representation theorem, there is a unique $Tf \in L^2(Y)$ such that

$$\int_{X\times Y} fg\,d\mu = \int_Y (Tf)g\,d\mu_Y \quad \text{for all } g \in L^2(Y). \tag{4}$$

This defines a map $T : L^2(X) \to L^2(Y)$. We need to check that T is a Markov operator.

Linearity of T follows from (4) and boundedness from (3). It is straightforward to show, by (4), that $f \ge 0$ implies $Tf \ge 0$ for all $f \in L^2(X)$. Moreover, for arbitrary $g \in L^2(Y)$, we see from (4) that

$$\int_Y (T1)g\,d\mu_Y = \int_{X\times Y} g\,d\mu = \int_Y g\,d\mu_Y.$$

Thus $T1 = 1$. It follows from the positivity of T that $T : L^\infty(X) \to L^\infty(Y)$. Finally, by taking $g = 1$ in (4), we have

$$\int_Y Tf\,d\mu_Y = \int_{X\times Y} f\,d\mu = \int_X f\,d\mu_X.$$

The relationship $\mu(A\times B)=\int_Y (T\chi_A)\chi_B\,d\mu_Y$ follows from (4) by setting $f=\chi_A$ and $g=\chi_B$. ■

COROLLARY 1. *There is a unique Markov operator* $T^*: L^\infty(Y)\to L^\infty(X)$ *such that*

$$\int_X f(T^*g)\,d\mu_X=\int_Y (Tf)g\,d\mu_Y \tag{5}$$

for all $f\in L^\infty(X)$ *and* $g\in L^\infty(Y)$.

The next most striking property of a Markov operator after its connection with μ is that Tf is a conditional expectation. To see this, it is convenient to introduce the projection map $\pi: X\times Y\to Y$.

COROLLARY 2. *If we regard* $f\in L^\infty(X,\mu_X)$ *as being a random variable on* $(X\times Y,\mu)$, *where* μ_X *and* μ_Y *are the marginals of* μ, *then*

$$(Tf)(y)=E(f\,|\,\pi=y)\qquad \mu_Y\text{-a.e.}$$

Proof. For every measurable $B\subseteq Y$, by equation (2) with $g=\chi_B$ in $L^\infty(Y)$ and $g=\chi_{\{\pi\in B\}}$ in $L^\infty(X\times Y)$, we have

$$\int_{\{\pi\in B\}} f\,d\mu=\int_B Tf\,d\mu_Y.$$

By the definition of conditional expectation, we get

$$\int_{\{\pi\in B\}} f\,d\mu=\int_B E(f\,|\,\pi=y)\,d\mu_Y(y). \tag{6}$$

The result is established. ■

We now develop a few other simple properties of Markov operators.

THEOREM 3. *Suppose that we are given a measurable map* $\phi: Y\to X$ *and a probability measure* μ_Y *on* Y. *Define a probability measure* μ_X *on* X *by* $\mu_X(A)=\mu_Y(\phi^{-1}(A))$ *for measurable* A *and a Markov operator* $T: L^\infty(X)\to L^\infty(Y)$ *by* $Tf=f\circ\phi$. *Then*:

1. *The mass of the probability measure* μ *induced on* $X\times Y$ *by* T *is concentrated on the graph of* ϕ.
2. μ *is an extremal element of the convex set*

$$\mathcal{M}=\{\nu:\nu \text{ is a probability measure on } X\times Y \text{ with } Y\text{-marginal } \mu_Y\}.$$

Proof. Let Φ be the graph of ϕ and let $\pi: X\times Y\to Y$ be the natural projection. For measurable $S\subseteq X\times Y$, define

$$\nu(S)=\mu_Y(\pi(S\cap\Phi)).$$

It is straightforward to show that ν is a measure on $X\times Y$. (It may be helpful to consider Figure 1.) Further, since $\nu(X\times Y)=\nu(\Phi)=\mu_Y(Y)=1$, we see that ν is a probability measure whose mass is concentrated on Φ. Next, for

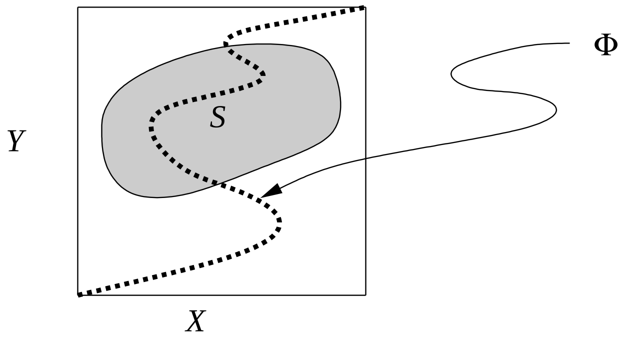

Fig. 1. S and the graph of ϕ

measurable $A \subseteq X$ and $B \subseteq Y$, it is easily checked that $\pi((A \times B) \cap \Phi) = \phi^{-1}(A) \cap B$. It follows that

$$\begin{aligned}\nu(A \times B) &= \mu_Y(\pi((A \times B) \cap \Phi)) = \mu_Y(\phi^{-1}(A) \cap B)\\ &= \int_Y \chi_{\phi^{-1}(A)} \chi_B \, d\mu_Y = \int_Y (\chi_A \circ \phi)\, \chi_B \, d\mu_Y\\ &= \int_Y (T\chi_A)\chi_B \, d\mu_Y = \mu(A \times B)\end{aligned}$$

where the last step is justified by Theorem 2. Since this holds for measurable sets of the form $A \times B$, we have $\mu = \nu$. Therefore the mass of μ is concentrated on Φ.

To show extremality of μ in $\mathcal{M}$, it suffices to suppose that we can write $\mu = \frac{1}{2}\xi + \frac{1}{2}\eta$ where $\xi, \eta \in \mathcal{M}$ and show that we must have $\mu = \xi = \eta$. Since the mass of μ is concentrated on Φ, the same must be true for ξ and η. For measurable $S \subseteq X \times Y$, it is easily shown that

$$S \cap \Phi = (X \times \pi(S \cap \Phi)) \cap \Phi.$$

Using this and the fact that ξ has Y-marginal μ_Y, we see that

$$\begin{aligned}\xi(S) &= \xi(S \cap \Phi) = \xi((X \times \pi(S \cap \Phi)) \cap \Phi) = \xi(X \times \pi(S \cap \Phi))\\ &= \mu_Y(\pi(S \cap \Phi)) = \mu(S)\end{aligned}$$

where the last step follows from the earlier part of this proof. Similarly, $\eta = \mu$. Thus μ is extremal in $\mathcal{M}$. ■

Next we consider another particularly simple Markov operator.

THEOREM 4. *Suppose $T : L^\infty(X) \to L^\infty(Y)$ is a Markov operator and μ is the associated probability measure on $X \times Y$. Then T has the form*

$$Tf = c_f, \qquad \textit{a constant associated with } f,$$

if and only if $\mu = \mu_X \times \mu_Y$. Further, $c_f = \int_X f \, d\mu_X = E(f)$, the expected value of f.

Proof. Suppose that $Tf = c_f$. First notice that

$$c_f = \int_Y c_f \, d\mu_Y = \int_Y Tf \, d\mu_Y = \int_X f \, d\mu_X = E(f).$$

Now let $A \subseteq X$ and $B \subseteq Y$ be measurable sets. Appealing to Theorem 2, we see that

$$\begin{aligned}\mu(A \times B) &= \int_{X\times Y} \chi_A(x)\chi_B(y)\, d\mu(x,y) = \int_B T\chi_A \, d\mu_Y \\ &= c_{\chi_A}\, \mu_Y(B) = \mu_X(A)\mu_Y(B).\end{aligned}$$

Hence $\mu = \mu_X \times \mu_Y$.

Now let us suppose that $\mu = \mu_X \times \mu_Y$. Let $f \in L^\infty(X)$ be a simple function, $f = \sum_{k=1}^n c_k \chi_{A_k}$, where each A_k is a measurable subset of X, and let B be a measurable subset of Y. Appealing to Theorem 2, we have

$$\begin{aligned}\int_B Tf \, d\mu_Y &= \int_{X\times Y} f(x)\chi_B(y)\, d\mu(x,y) = \sum_k c_k \int_{X\times Y} \chi_{A_k \times B}\, d\mu \\ &= \sum_k c_k \mu(A_k \times B) = \sum_k c_k \mu_X(A_k)\mu_Y(B) = \int_B E(f)\, d\mu_Y.\end{aligned}$$

Thus $Tf = E(f)$.

Now let f be an arbitrary element of $L^\infty(X)$. There exists a sequence of simple functions $\{f_n\}$ such that $f_n \to f$ in $L^\infty(X)$. Since T is bounded, $Tf_n \to Tf$ in $L^\infty(Y)$, and since each Tf_n is a constant, the same must be true for Tf. ∎

We now describe a method, mentioned in [7], of constructing a Markov operator. The proof is easy and is essentially that of [7].

Theorem 5 (Kulpa). *Let (X, μ_X) and (Y, μ_Y) be probability spaces and K be a nonnegative, real-valued, measurable function on $X \times Y$ such that*

$$\int_X K(x,y)\, d\mu_X(x) = 1 \quad \mu_Y\text{-a.e.}, \quad \text{and} \quad \int_Y K(x,y)\, d\mu_Y(y) = 1 \quad \mu_X\text{-a.e.}$$

Then:

1. *The equation*

   $$Tf(y) = \int_X K(x,y) f(x)\, d\mu_X(x)$$

   defines a Markov operator from $L^\infty(X)$ to $L^\infty(Y)$.
2. *The probability measure μ associated with T satisfies*

   $$\mu(S) = \int_S K \, d(\mu_X \times \mu_Y)$$

   for measurable $S \subseteq X \times Y$.

3. *The adjoint operator* $T^* : L^\infty(Y) \to L^\infty(X)$ *is given by*

$$T^*g(x) = \int_Y K(x,y)g(y)\,d\mu_Y(y).$$

REMARK 1. We call K the *kernel* of the Markov operator T or of μ with respect to the marginals μ_X and μ_Y.

COROLLARY 3. *Let μ be a probability measure on $X \times Y$ with marginals μ_X and μ_Y. Then the Markov operator $T : L^\infty(X) \to L^\infty(Y)$ has a kernel $K : X \times Y \to \mathbb{R}$ if and only if μ is absolutely continuous with respect to $\mu_X \times \mu_Y$. In this case,*

$$K = \frac{d\mu}{d(\mu_X \times \mu_Y)}.$$

Proof. We only consider the direction in which we start with $\mu \ll \mu_X \times \mu_Y$. Then by the Radon–Nikodym theorem, μ has a Radon–Nikodym derivative with respect to $\mu_X \times \mu_Y$. Let us set K equal to this Radon–Nikodym derivative. From this and $\mu(A \times B) = \int_B T\chi_A\,d\mu_Y$ one can deduce

$$\int_B Tf\,d\mu_Y = \int_B \int_X K(x,y)f(x)\,d\mu_X(x)\,d\mu_Y(y)$$

whenever $f \in L^\infty(X,\mu_X)$ and B is a measurable subset of Y. It is then easily seen that

$$\int_X K(x,y)\,d\mu_Y(y) = 1 \quad \text{and} \quad \int_Y K(x,y)\,d\mu_X(x) = 1. \blacksquare$$

3. Copulas and Markov operators. Let us recall that an n-copula, $n \geq 2$, is a function $C : I^n \to \mathbb{R}$ satisfying the following:

1. $C(x_1,\dots,x_n) = 0$ if any $x_k = 0$;
2. $C(x_1,\dots,x_n) = x_k$ if $x_j = 1$ for all $j \neq k$;
3. C is n-increasing. (This condition is defined in [13] and [15] in terms of certain inequalities. It is more convenient for our purposes to use the equivalent condition that there is a measure μ on I^n such that

$$C(x_1,\dots,x_n) = \mu([0,x_1] \times \dots \times [0,x_n])$$

for all $(x_1,\dots,x_n) \in I^n$.)

We call the first two conditions the *boundary conditions* on a copula. We refer to μ as the probability measure associated with C. Notice that if we factor I^n into one-dimensional and $(n-1)$-factors, say, $I^n = I^{(n-1)} \times I$, then the marginal measures on I must always be λ, one-dimensional Lebesgue measure.

Another way to define an n-copula C is to say that it is a joint distribution function for an n-tuple of random variables $(X_1,\dots,X_n)$ such that each X_k is uniformly distributed over I.

Every n-copula C with its associated probability measure μ and every factorization $I^n = I^p \times I^q$, where $p, q \geq 1$, gives rise to an associated Markov operator $T : L^\infty(I^p) \to L^\infty(I^q)$.

Example 3. One important n-copula is

$$\Pi^n(x_1, \ldots, x_n) = x_1 \cdots x_n.$$

The associated probability measure is λ^n, n-dimensional Lebesgue measure. The Markov operator $T : L^\infty(I^p) \to L^\infty(I^q)$ associated with Π^n is

$$Tf(y) = \int_{I^p} f(x)\, dx \qquad \lambda^q\text{-a.e.}$$

where $x \in I^p$ and $y \in I^q$.

Example 4. Another important n-copula is

$$M^n(x_1, \ldots, x_n) = \min(x_1, \ldots, x_n).$$

The mass of the associated probability measure μ is distributed uniformly along the "diagonal" $x_1 = \cdots = x_n$ in I^n. Similarly, the mass of each marginal measure μ_p and μ_q is uniformly distributed along the "diagonal" $x_1 = \cdots = x_p$ and $x_{p+1} = \cdots = x_n$ of I^p and I^q respectively. It follows that $f \in L^\infty(I^p, \mu_p)$ is fully determined by considering $f(t^p)$ where t^p is the p-tuple $(t, \ldots, t) \in I^p$ for every $t \in I$. Then the Markov operator $T : L^\infty(I^p) \to L^\infty(I^q)$ is defined by

$$Tf(t^q) = f(t^p) \qquad \mu_q\text{-a.e.}$$

3.1. *Differentiable n-copulas.* In what follows, we will find it convenient to refer to densities of copulas: For any n-copula C, we use the notation

$$C^{(n)} = \frac{\partial^n C}{\partial x_1 \cdots \partial x_n},$$

assuming the mixed partial exists. We say that $C^{(n)}$ is a *density* for C provided

1. $C^{(n)}$ exists λ^n-a.e. on I^n.
2. For all Borel subsets S of I^n, we have

$$\mu(S) = \int_S C^{(n)}\, d\lambda^n$$

   where μ is the probability measure associated with C.

Let $I^n = I^p \times I^q$, $p, q \geq 1$, and let η and ξ be the marginal measures of μ on I^p and I^q respectively. Set $1^p = (1, \ldots, 1) \in I^p$ and $1^q = (1, \ldots, 1) \in I^q$. Assuming C has a density $C^{(n)}$, it will also be useful for us to set

$$C^{(p)}(x, y) = \frac{\partial^p C}{\partial x_1 \cdots \partial x_p}(x, y), \qquad C^{(q)}(x, y) = \frac{\partial^q C}{\partial x_{p+1} \cdots \partial x_n}(x, y)$$

where $x = (x_1, \ldots, x_p) \in I^p$ and $y = (x_{p+1}, \ldots, x_n) \in I^q$. From the fact that $\eta([0,x]) = \mu([0,x] \times I^q) = C(x, 1^q)$, we deduce

$$\eta(A) = \int_A C^{(p)}(x, 1^q)\, dx.$$

Similarly,

$$\xi(B) = \int_B C^{(q)}(1^p, y)\, dy.$$

Thus

$$(\eta \times \xi)(S) = \int_S C^{(p)}(x, 1^q) C^{(q)}(1^p, y)\, dx\, dy$$

for S a Borel subset of I^n. Using this notation, we then have the following:

THEOREM 6. *Let C be an n-copula with continuous density $C^{(n)}$. If T is the Markov operator $T : L^\infty(I^p, \eta) \to L^\infty(I^q, \xi)$, then T has the kernel*

$$K(x, y) = \frac{C^{(n)}(x, y)}{C^{(p)}(x, 1^q)\, C^{(q)}(1^p, y)} \qquad \eta \times \xi\text{-a.e.}$$

Further,

$$Tf(y) = \int_{I^p} \frac{C^{(n)}(x, y)}{C^{(q)}(1^p, y)} f(x)\, dx \qquad \xi\text{-a.e.}$$

for $f \in L^\infty(I^p, \eta)$ where it may be assumed, without loss of generality, that $f(x) = 0$ whenever $C^{(p)}(x, 1^q) = 0$.

Proof. If the kernel of T exists, it has the form

$$K = \frac{d\mu}{d(\eta \times \xi)},$$

so showing the existence of K requires us to show that $\mu \ll \eta \times \xi$. Once we have done that, it is trivial to check the forms of $K(x, y)$ and $Tf(y)$ using the integral formulas for μ, η, ξ, and $\eta \times \xi$.

Choose $(x, y) \in I^p \times I^q$ such that $C^{(n)}(x, y) > 0$. There exist $\varepsilon > 0$ and open sets J_p and J_q in I^p and I^q respectively such that $x \in J_p$, $y \in J_q$, and $C^{(n)}(u, v) \geq \varepsilon$ for all $u \in J_p$ and $v \in J_q$. Then

$$C^{(q)}(1^p, y) = \int_{I^p} C^{(n)}(u, y)\, du \geq \varepsilon \lambda^p(J_p) > 0.$$

Similarly, $C^{(p)}(x, 1^p) > 0$. Thus $C^{(p)}(x, 1^q)\, C^{(q)}(1^p, y) > 0$.

Choose a Borel subset $S \subseteq I^n$ such that $(\eta \times \xi)(S) = 0$. Set

$$S_0 = \{(x, y) \in S : C^{(p)}(x, 1^q) C^{(q)}(1^p, y) = 0\}$$

and $S_1 = S - S_0$. We know that if $C^{(p)}(x, 1^q)C^{(q)}(1^p, y) = 0$, then $C^{(n)}(x, y) = 0$. So

$$\begin{aligned}\mu(S) &= \mu(S_1) + \int_{S_0} C^{(n)}\, d\lambda^n = \mu(S_1)\\ &= \int_{S_1} \frac{C^{(n)}(x,y)}{C^{(p)}(x,1^q)C^{(q)}(1^p,y)}\, C^{(p)}(x,1^q)C^{(q)}(1^p,y)\, dx\, dy\\ &= \int_{S_1} K\, d(\eta\times\xi) = 0.\end{aligned}$$

Thus $\mu \ll \eta \times \xi$. ■

REMARK 2. If $p = n - 1$ and $q = 1$, then $C(1^{n-1}, y) = y$ so that $C^{(1)}(1^{n-1}, y) = 1$. Then T takes on the particularly simple form of

$$Tf(y) = \int_{I^{n-1}} \frac{\partial^n C}{\partial x_1 \cdots \partial x_n}(x, y)\, f(x)\, dx \qquad \xi\text{-a.e.}$$

where $x \in I^{n-1}$ and $y \in I$.

3.2. *First partials of copulas.* Higher order partial derivatives of copulas may fail to exist or fail to provide useful information about the copula. For example, if $M^2(x, y) = \min(x, y)$, then the mixed second order partial of M^2 is zero everywhere except along the diagonal $x = y$ where it is undefined.

However, first partials exist almost everywhere and the copula can be reconstructed from them.

If we factor I^n thus: $I^n = I^p \times I \times I^q$, where the indicated I is the kth factor of I^n, then we designate

$$I^{(n,k)} = I^p \times I^q = \{(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) : (x_1, \dots, x_n) \in I^n\}.$$

This may merely seem a way to write I^{n-1}, but we are to think that the x_j's in $I^{(n,k)}$ "remember" their positions in I^n. By a slight abuse of notation, we also feel free to interpret $I^{(n,k)} \times I$ as $I^p \times I \times I^q$ rather than $I^p \times I^q \times I$.

Let μ be the probability measure associated with the n-copula C, $n \geq 2$. We denote the marginal measure on $I^{(n,k)}$ as $\mu^{(n,k)}$. So $\mu^{(n,k)}(S) = \mu(S \times I)$ when S is a Borel subset of $I^{(n,k)}$. For $k = 1, \dots, n$, we have a Markov operator

$$T^k : L^\infty(I^{(n,k)}) \to L^\infty(I, \lambda)$$

associated with μ.

It can also be useful to keep in mind that if $X_k : I^n \to I$ is the projection map $X_k(t_1, \dots, t_n) = t_k$, then

$$T^k f(x_k) = E(f \mid X_k = x_k).$$

Here is the connection between first partials of copulas and Markov operators:

Theorem 7. *Let C be an n-copula with associated probability measure μ. Then for $i = 1, \ldots, n$ and $(x_1, \ldots, t, \ldots, x_n) \in I^n$ we have*

$$\begin{aligned}\frac{\partial C}{\partial x_i}(x_1, \ldots, t, \ldots, x_n) &= \mu([0, x_1] \times \cdots I \cdots \times [0, x_n] \mid X_i = t)\\ &= T^i \chi_R(t) \quad \lambda(t)\text{-a.e.}\end{aligned}$$

where t and I occur in the ith position and

$$R = [0, x_1] \times [0, x_{i-1}] \times [0, x_{i+1}] \times \cdots \times [0, x_n].$$

Proof. We consider only the case $i = 1$, that is, differentiation with respect to the first variable. Fix $(x_2, \ldots, x_n)$ and consider $0 < t < 1$. Since copulas are monotonic in each variable, we see that $\frac{\partial C}{\partial x_1}(t, x_2, \ldots, x_n)$ exists $\lambda(t)$-a.e. Copulas satisfy a Lipschitz condition, hence are absolutely continuous when considered as a function of t. It follows that

$$\int_0^{x_1} \frac{\partial C}{\partial x_1}(t, x_2, \ldots, x_n)\, dt = C(x_1, x_2, \ldots, x_n) = \mu([0, x_1] \times \cdots \times [0, x_n]).$$

We see from this that

$$\begin{aligned}\frac{\partial C}{\partial x_1}(t, x_2, \ldots, x_n) &= \mu(I \times [0, x_2] \times \cdots \times [0, x_n] \mid X_1 = t)\\ &= T^1 \chi_{[0,x_2]\times\cdots\times[0,x_n]}(t) \quad \lambda(t)\text{-a.e.} \;\blacksquare\end{aligned}$$

3.3. *Products of copulas*

3.3.1. *Composition of Markov operators.* A product of copulas, chiefly 2-copulas, is introduced and studied in [5]. It is shown there that this product has an interesting connection with Markov processes. We use the notion of Markov operators to extend that definition. We show that one can "multiply" a $(p+q)$-copula and $(q+r)$-copula to produce a $(p+r)$-copula provided the q-marginals of the factors are the same.

Suppose that A and B are $(p + q)$- and $(q + r)$-copulas respectively where $p, q, r \geq 1$. Let μ and ν be the probability measures on I^{p+q} and I^{q+r} associated with A and B respectively, and let μ_p, μ_q be the marginals of μ on I^p and I^q and ν_q, ν_r be the marginals of ν on I^q and I^r. It is also convenient at this point to introduce the notation 1^p for the p-tuple $(1, \ldots, 1)$.

Theorem 8. *Let $T_A : L^\infty(I^p, \mu_p) \to L^\infty(I^q, \mu_q)$ and $T_B : L^\infty(I^q, \nu_q) \to L^\infty(I^r, \nu_r)$ be the Markov operators associated with μ and ν respectively. If $\mu_q = \nu_q$, let $T = T_B T_A : L^\infty(I^p, \mu_p) \to L^\infty(I^r, \nu_r)$ be the Markov operator with associated probability measure ξ on I^{p+r}. Then ξ is generated by a $(p + r)$-copula which we designate AB and which satisfies*

$$\begin{aligned}AB(x_1, \ldots, x_p, 1^r) &= A(x_1, \ldots, x_p, 1^q),\\ AB(1^p, x_{p+1}, \ldots, x_{p+r}) &= B(1^q, x_{p+1}, \ldots, x_{p+r}).\end{aligned}$$

Further, if $q = 1$, then $\mu_q = \nu_q = \lambda$ and

$$(7)\qquad AB(x_1, \ldots, x_{p+r}) = \int_0^1 \frac{\partial A}{\partial x_{p+1}}(x_1, \ldots, x_p, t)\, \frac{\partial B}{\partial x_1}(t, x_{p+1}, \ldots, x_{p+r})\, dt.$$

Proof. The measure ξ is associated with T by the requirement that

$$\xi(R \times S) = \int_S T\chi_R\, d\mu_p$$

when R and S are Borel sets of I^p and I^r respectively. Then AB is defined by

$$AB(x) = \xi([0, x])$$

where $x = (x_1, \ldots, x_{p+r}) \in I^{p+r}$ and $[0, x] = [0, x_1] \times \cdots \times [0, x_{p+r}]$.

The only condition we need to check to show that AB is a copula is

$$AB(1, \ldots, 1, x_k, 1, \ldots, 1) = x_k,$$

that is, we need to show that ξ has uniform 1-marginals. Consider the case where K is a Borel set of I, $R = K \times I^{p-1}$, and $S = I^r$. Applying the definition of Markov operator, we have

$$\xi(K \times I^{p-1} \times I^r) = \int_{I^r} T_B T_A \chi_R\, d\nu_r = \int_{I^q} T_A \chi_R\, d\nu_q = \int_{I^p} \chi_R\, d\mu_p = \lambda(K).$$

If, on the other hand, we have $R = I^p$ and $S = I^{r-1} \times K$, then

$$\xi(I^p \times I^{r-1} \times K) = \int_{I^{r-1} \times K} T_B T_A 1\, d\nu_r = \nu_r(I^{r-1} \times K) = \lambda(K).$$

Thus AB is a copula.

Next let $(x_1, \ldots, x_p) \in I^p$ and set $K = [0, x_1] \times \cdots \times [0, x_p]$. Then

$$\begin{aligned} AB(x_1, \ldots, x_p, 1^r) = \xi(K \times I^r) &= \int_{I^r} T\chi_K\, d\nu_r = \int_{I^p} \chi_K\, d\mu_p \\ &= \mu(K \times I^q) = A(x_1, \ldots, x_p, 1^q). \end{aligned}$$

The proof that $AB(1^p, x_{p+1}, \ldots, x_{p+r}) = B(1^q, x_{p+1}, \ldots, x_{p+r})$ is similar. ∎

Remark 3. The notation AB is deficient in the following sense: We may think of A as a function of $(x_1, \ldots, x_p) \in I^p$ and $(y_1, \ldots, y_q) \in I^q$ and write

$$A(x_1, \ldots, x_p, y_1, \ldots, y_q).$$

In a similar fashion, we may write

$$B(y_1, \ldots, y_q, z_1, \ldots, z_r)$$

where $(z_1, \ldots, z_r) \in I^r$. It is then natural to write

$$AB(x_1, \ldots, x_p, z_1, \ldots, z_r).$$

The variables $y_1, \ldots, y_q$ have been eliminated in forming AB. The deficiency lies in the fact that there are many ways to choose $y_1, \ldots, y_q$; there are many possibilities for q and there is no necessity to choose them all from the "front end" or "rear end" of the domains. Hence there are many products that can be formed. In this particular discussion, for example, it might be more useful to use a symbol such as

$$A \underset{y_1,\ldots,y_q}{*} B$$

for AB.

3.3.2. *A minimal "join" of two copulas.* The "product" that we define in equation (8) below is clearly one suggested by our considerations of Markov operators though not directly defined by them. This construction was originally introduced in [5] and used to study Markov processes. Here, however, we investigate a different question.

When given p- and q-copulas A and B respectively, we may wish to find an n-copula C having A and B as marginals. A simple and easy way to do this is to set $C(x,y) = A(x)B(y)$ where $x \in I^p$ and $y \in I^q$. Is there any less obvious way to do this? Is there some way to do this that "crowds A and B as close together" as possible?

We offer one answer to this question. In what follows, it is intended that

$$x = (x_2,\ldots,x_p) \in I^{p-1}, \quad y = (y_2,\ldots,y_q) \in I^{q-1}, \quad t,u \in I.$$

THEOREM 9. *Let A and B be p- and q-copulas respectively. Define $C : I^{p+q-1} \to \mathbb{R}$ by*

$$C(u,x,y) = \int_0^u \frac{\partial A}{\partial t}(t,x)\,\frac{\partial B}{\partial t}(t,y)\,dt. \tag{8}$$

Then C is a $(p+q-1)$-copula having A and B as marginals. Further, $p+q-1$ is the minimal dimension for which one can, in general, construct a copula C having given p- and q-copulas A and B as marginals.

Proof. The fact that C is a copula is noted in [5]. It is easily checked that A and B are marginals of C.

To prove the minimality of $p+q-1$, set $A(x_1,\ldots,x_p) = \min(x_1,\ldots,x_p)$ and $B(y_1,\ldots,y_q) = y_1\cdots y_q$. We may suppose that C is an n-copula having A and B as marginals and such that $n \le p+q-1$. It will suffice to show that we must have $n = p+q-1$. We may, without loss of generality, suppose that

$$C(x_1,\ldots,x_p,1^{n-p}) = A(x_1,\ldots,x_p).$$

Because A is the copula min, we see that the support of μ, the probability measure associated with C, must lie in

$$P = \{(t,\ldots,t,u_1,u_2,\ldots,u_{n-p}) : t,u_k \in I \text{ for } k = 1,\ldots,n-p\} \subseteq I^n.$$

Now let us suppose that $n < p+q-1$. Choose nondegenerate intervals J_1, J_2 in I such that $J_1 \cap J_2 = \emptyset$. Set

$$K = I^{p-2} \times J_1 \times J_2 \times I^{n-p}.$$

We may, again without loss of generality, assume that B occurs as a marginal of C in such a way that

$$\mu(K) = \mu_B(J_1 \times J_2 \times I^{q-2}) = \lambda^2(J_1 \times J_2) > 0.$$

However, we also have $K \cap P = \emptyset$, so $\mu(K) = 0$. From this contradiction, we conclude that $n = p + q - 1$. ∎

3.4. *Sequential convergence.* In the next two definitions, let $C, C_1, C_2, \dots$ be n-copulas and $\mu, \mu_1, \mu_2, \dots$ the respective associated probability measures on I^n. Recall the definition of $I^{(n,i)}$ from Subsection 3.2, and let ν_i be the marginal measure of μ on $I^{(n,i)}$ and ν_{ki} be the marginal measure of μ_k on $I^{(n,i)}$.

Definition 2. We write

$$C_k \xrightarrow{\mathfrak{M}} C$$

provided that for all $i \in \{1, \dots, n\}$ and for all bounded, measurable $f : I^{n-1} \to \mathbb{R}$ we have

$$\lim_{k\to\infty} \int_I |T_k^i f(t) - T^i f(t)|\, dt = 0$$

where $T^i : L^\infty(I^{(n,i)}, \nu_i) \to L^\infty(I, \lambda)$ and $T_k^i : L^\infty(I^{(n,i)}, \nu_{ki}) \to L^\infty(I, \lambda)$ are the Markov operators associated with μ and μ_k respectively.

Notice in this last definition that although a Markov operator is applied to an equivalence class of functions, we choose f to be an actual function. If we were, for example, to denote the equivalence class of f in $L^\infty(I^{(n,i)}, \nu_i)$ as $[f]_{\nu_i}$, then for a given bounded, measurable f, we might more carefully write $T_k^i f - T^i f$ as $T_k^i([f]_{\nu_{ki}}) - T^i([f]_{\nu_i})$.

Definition 3. We write

$$C_k \xrightarrow{\partial} C$$

provided that for all $i \in \{1, \dots, n\}$ we have

$$\lim_{k\to\infty} \int_I \left| \frac{\partial C_k}{\partial x_i}(x_1, \dots, t, \dots, x_n) - \frac{\partial C}{\partial x_i}(x_1, \dots, t, \dots, x_n) \right| dt = 0$$

where t is the ith coordinate of the n-tuple.

Theorem 10. *Let $C, C_1, C_2, \dots$ be n-copulas and consider the following conditions*:

(1) $C_k \xrightarrow{\mathfrak{M}} C$.
(2) $C_k \xrightarrow{\partial} C$.
(3) $C_k \to C$ *uniformly.*
Then (1)⇒(2)⇒(3).

Proof. (1)⇒(2). If we set $f = \chi_{[0,x_2]\times\cdots\times[0,x_n]}$ in

$$\lim_{k\to\infty} \int_I |T_k^1 f - T^1 f|\, d\lambda = 0,$$

the desired result follows immediately.

(2)⇒(3). Since

$$\begin{aligned}|C_k(x_1,x_2,\ldots,x_n) - C(x_1,x_2,\ldots,x_n)|\\ &= \left|\int_0^{x_1}\left(\frac{\partial C_k}{\partial x_1}(t,x_2,\ldots,x_n) - \frac{\partial C}{\partial x_1}(t,x_2,\ldots,x_n)\right)dt\right|\\ &\le \int_0^1 \left|\frac{\partial C_k}{\partial x_1}(t,x_2,\ldots,x_n) - \frac{\partial C}{\partial x_1}(t,x_2,\ldots,x_n)\right| dt,\end{aligned}$$

the result follows. ■

We want to give examples or establish conditions under which these convergences might occur. We begin with some "natural" approximations of copulas that are again copulas.

It is useful to first describe a partition of I^n into cells: Let m be a "large" natural number and define

$$I^n_{m,i} = \left[\frac{i_1}{m}, \frac{i_1+1}{m}\right] \times \cdots \times \left[\frac{i_n}{m}, \frac{i_n+1}{m}\right]$$

where i is the multi-index $(i_1,\ldots,i_n)$ with each i_k in $\{0,1,\ldots,m-1\}$.

Now here are two approximations of copulas that are investigated in [8], [9], and [11]:

EXAMPLE 5 (Checkerboard approximation). Let C be a given n-copula and μ the associated probability measure. For $i = (i_1,\ldots,i_n)$, let χ_i be the characteristic function of $I^n_{m,i}$. Then the $m\times\cdots\times m$ *checkerboard approximation* C_m of C is constructed by replacing μ with a probability measure μ_m on I^n in such a way that μ_m has constant density $\mu(I^n_{m,i})/\lambda^n(I^n_{m,i}) = m^n\mu(I^n_{m,i})$ on each cell $I^n_{m,i}$. Of course, C_m is the n-copula associated with μ_m. If we permit i/m to stand for the point $(i_1/m,\ldots,i_n/m) \in I^n$, then we have $C_m(i/m) = C(i/m)$ for all i, and the density of C_m is given by

$$C_m^{(n)} = \frac{\partial^n C_m}{\partial x_1\cdots\partial x_n} = m^n \sum_i \mu(I^n_{m,i})\chi_i. \tag{9}$$

Example 6 (Bernstein approximation). It turns out that the standard Bernstein approximation of a copula is again a copula. (See [11] for more detail on this example.) Let C and μ be as before.

We describe the Bernstein approximation process for C: The mth degree Bernstein polynomial $b_{i,m} : I \to \mathbb{R}$ is given by

$$b_{i,m}(t) = \binom{m}{i} t^i (1-t)^{m-i}, \qquad i = 0, 1, \ldots, m.$$

We extend this to $B^n_{i,m} : I^n \to \mathbb{R}$ by taking i to be a multi-index, $i = (i_1, \ldots, i_n)$, with each i_k in $\{0, 1, \ldots, m\}$, and setting

$$B^n_{i,m}(x) = b_{i_1,m}(x_1) b_{i_2,m}(x_2) \cdots b_{i_n,m}(x_n)$$

where $x = (x_1, \ldots, x_n) \in I^n$. Then by the $m \times \cdots \times m$ *Bernstein approximation* to C we mean

$$A_m(x) = \sum_i C(i/m) B^n_{i,m}(x)$$

where $x \in I^n$ and i is a multi-index.

We accept that A_m is an n-copula, and since C is continuous, it is a standard result of analysis that A_m converges uniformly to C on I^n. If we extend our definition of $b_{i,m}$ by setting

$$b_{-1,m} = b_{m+1,m} = 0,$$

then we can show that

$$b'_{i,m} = m\,(b_{i-1,m-1} - b_{i,m-1}), \qquad i = 0, 1, \ldots, m.$$

It is then possible to prove that the density of A_m is given by

$$A^{(n)}_m = \frac{\partial^n A_m}{\partial x_1 \cdots \partial x_n}(x_1, \ldots, x_n) = m^n \sum_i \mu(I^n_{m,i}) B^n_{i,m-1}. \tag{10}$$

Coming back to the theme of convergence, we have the following result from [11]:

Theorem 11. *Let C_m and A_m be, respectively, the $m \times \cdots \times m$ checkerboard and Bernstein approximations to a given n-copula C. Then $C_m, A_m \xrightarrow{\partial} C$ as $m \to \infty$, however we do not, in general, have either $C_m \xrightarrow{\mathfrak{M}} C$ or $A_m \xrightarrow{\mathfrak{M}} C$.*

$\mathfrak{M}$-convergence seems to require much stronger conditions than ∂-convergence or uniform convergence. We do, however, establish some conditions under which it does occur.

Theorem 12. *Let $C, C_1, C_2, \ldots$ be n-copulas with associated probability measures $\mu, \mu_1, \mu_2, \ldots$ respectively. Assume that*

1. *C and each C_k has a density $C^{(n)}$ and $C^{(n)}_k$ respectively,*

2. $C_k^{(n)} \to C^{(n)}$ *in* $L^1(I^n, \lambda^n)$.

Then $C_k \xrightarrow{\mathfrak{M}} C$.

Proof. It suffices to consider the factorization $I^n = I \times I^{n-1}$ of the unit cube and the Markov operators $T, T_k : L^\infty(I^{n-1}) \to L^\infty(I)$ associated with μ and μ_k respectively. We know by Remark 2 that

$$Tf(x) = \int\limits_{I^{n-1}} C^{(n)}(x,y) f(y)\, dy \quad \text{and} \quad T_k f(x) = \int\limits_{I^{n-1}} C_k^{(n)}(x,y) f(y)\, dy$$

where $x \in I$ and $y \in I^{n-1}$. Then for bounded, measurable $f : I^{n-1} \to \mathbb{R}$, we have

$$\begin{aligned} 0 \le \int\limits_I |T_k f - Tf|\, d\lambda &= \int\limits_I \Big| \int\limits_{I^{n-1}} (C_k^{(n)} - C^{(n)})(x,y) f(y)\, dy \Big|\, dx \\ &\le \int\limits_{I^n} |(C_k^{(n)} - C^{(n)})(x,y)| \|f\|_\infty \, dy\, dx \to 0 \quad \text{as } k \to \infty. \end{aligned}$$

Hence $C_k \xrightarrow{\mathfrak{M}} C$. ∎

An obvious next step is to ask for $C, C_1, C_2, \ldots$ that satisfy the hypotheses of Theorem 12. Toward that end, we assume in the next two lemmas that C is a given n-copula with associated probability measure μ and density $C^{(n)}$.

LEMMA 1. *If* $C^{(n)}$ *is continuous on* I^n *and* C_m *is the* $m \times \cdots \times m$ *checkerboard approximation of* C, *then* $C_m^{(n)} \to C^{(n)}$ *uniformly* λ^n-*a.e. on* I^n *as* $m \to \infty$.

Proof. Choose $\varepsilon > 0$. By the uniform continuity of $C^{(n)}$ on I^n, we can choose m so large that if $y, z \in I^n_{m,i}$, then $|C^{(n)}(y) - C^{(n)}(z)| < \varepsilon$. Notice that for every multi-index $i = (i_1, \ldots, i_n)$ there exists $y_i \in I^n_{m,i}$ such that

$$\mu(I^n_{m,i}) = \int\limits_{I^n_{m,i}} C^{(n)}\, d\lambda^n = C^{(n)}(y_i)\, \frac{1}{m^n}.$$

Let us choose $x \in I^n$ such that x lies in the interior of some cell $I^n_{m,j}$; the set of such x has λ^n-measure 1. Recalling the formula (9) and the fact that χ_i is the characteristic function of $I^n_{m,i}$, we have

$$\begin{aligned} |C_m^{(n)}(x) - C^{(n)}(x)| &= \Big| m^n \sum_i \mu(I^n_{m,i}) \chi_i(x) - C^{(n)}(x) \Big| \\ &= |C^{(n)}(y_j) - C^{(n)}(x)| < \varepsilon. \ \blacksquare \end{aligned}$$

LEMMA 2. *If* $C^{(n)}$ *is continuous on* I^n *and* A_m *is the* $m \times \cdots \times m$ *Bernstein approximation of* C, *then* $A_m^{(n)} \to C^{(n)}$ *uniformly on* I^n *as* $m \to \infty$.

Proof. Choose $\varepsilon > 0$. By the uniform continuity of $C^{(n)}$, we can choose m so large that if $|y-z| < n/(m-1)$, then $|C^{(n)}(y) - C^{(n)}(z)| < \varepsilon$. As in the last lemma, for every multi-index $i = (i_1, \dots, i_n)$, there exists $y_i \in I^n_{m,i}$ such that $\mu(I^n_{m,i}) = C^{(n)}(y_i)/m^n$. Recalling that i/m stands for a point in $I^n_{m,i}$, we see that $|y_i - i/m| < n/m$ so that $|C^{(n)}(y_i) - C^{(n)}(i/m)| < \varepsilon$. We shall also make use of the fact that $|i/m - i/(m-1)| < n/(m-1)$ so that

$$\left| C^{(n)}(y_i) - C^{(n)}\left(\frac{i}{m-1}\right)\right| < 2\varepsilon.$$

In what follows, we use the notation of Example 6. We know that $\{b_{i,m}\}_{i=0}^m$ is a partition of unity for I, so it follows that $\{B^n_{i,m}\}_i$ is a partition of unity for I^n where i is now a multi-index. Recalling the expression in (10) for $A^{(n)}_m$, we see that

$$\begin{aligned}
|A^{(n)}_m - C^{(n)}| &= \left| m^n \sum_i \mu(I^n_{m,i}) B^n_{i,m} - C^{(n)} \right| = \left| \sum_i C^{(n)}(y_i) B^n_{i,m} - C^{(n)} \right| \\
&\le \left| \sum_i C^{(n)}\left(\frac{i}{m-1}\right) B^n_{i,m} - C^{(n)} \right| \\
&\quad + \left| \sum_i \left(C^{(n)}(y_i) - C^{(n)}\left(\frac{i}{m-1}\right)\right) B^n_{i,m} \right| \\
&\le \left| \sum_i C^{(n)}\left(\frac{i}{m-1}\right) B^n_{i,m} - C^{(n)} \right| + 2\varepsilon.
\end{aligned}$$

Now $\sum_i C^{(n)}(i/(m-1))\, B^n_{i,m}$ is the $(m-1)\times\cdots\times(m-1)$ Bernstein approximation to $C^{(n)}$ and must converge uniformly to $C^{(n)}$ since $C^{(n)}$ is continuous. This establishes the result. ∎

Uniform convergence λ^n-a.e. implies L^1-convergence, so by Theorem 12, we have the following:

THEOREM 13. *If C is an n-copula such that $C^{(n)}$ is continuous and if C_m and A_m are the checkerboard and Bernstein approximations of C respectively, then $C_m, A_m \xrightarrow{\mathfrak{M}} C$.*

Acknowledgements. We are grateful to the referee for his suggestions and corrections that have improved this paper.

References

[1] R. B. Ash, *Real Analysis and Probability*, Academic Press, Orlando, FL, 1970.

[2] J. Behboodian, A. Dolati and M. Úbeda-Flores, *A multivariate version of Gini's rank association coefficient*, Statist. Papers 48 (2007), 295–304.

[3] J. R. Brown, *Approximation theorems for Markov operators*, Pacific J. Math. 16 (1966), 13–23.
[4] H. Carley and M. D. Taylor, *A new proof of Sklar's theorem*, in: Distributions with Given Marginals and Statistical Modelling, C. M. Cuadras, J. Fortiana and J. A. Rodriguez-Lallena (eds.), Kluwer, Dordrecht, 2002, 29–34.
[5] W. F. Darsow, B. Nguyen, and E. T. Olsen, *Copulas and Markov processes*, Illinois J. Math. 36 (1992), 600–642.
[6] A. Dolati and M. Úbeda-Flores, *On measures of multivariate concordance*, J. Probab. Statist. Sci. 4 (2006), 147–163.
[7] T. Kulpa, *On approximation of copulas*, Int. J. Math. Math. Sci. 22 (1999), 259–269.
[8] X. Li, P. Mikusiński, H. Sherwood and M. D. Taylor, *On approximation of copulas*, in: Distributions with Given Marginals and Moment Problems, V. Beneš and J. Štěpán (eds.), Kluwer, Dordrecht, 1997, 107–116.
[9] X. Li, P. Mikusiński, and M. D. Taylor, *Strong approximation of copulas*, J. Math. Anal. Appl. 225 (1998), 608–623.
[10] —, —, —, *Remarks on convergence of Markov operators*, Houston J. Math. 28 (2002), 907–916.
[11] P. Mikusiński and M. D. Taylor, *Some approximations of n-copulas*, in preparation.
[12] R. B. Nelsen, *Concordance and copulas: A survey*, in: Distributions with Given Marginals and Statistical Modelling, C. M. Cuadras, J. Fortiana and J. A. Rodriguez-Lallena (eds.), Kluwer, Dordrecht, 2002, 169–177.
[13] —, *An Introduction to Copulas*, 2nd ed., Springer, New York, 2006.
[14] H. L. Royden, *Real Analysis*, 3rd ed., Prentice-Hall, Englewood Cliffs, NJ, 1988.
[15] B. Schweizer and A. Sklar, *Probabilistic Metric Spaces*, Dover, Mineola, NY, 2005.
[16] M. D. Taylor, *Multivariate measures of concordance*, Ann. Inst. Statist. Math. 59 (2007), 789–806.

Department of Mathematics
University of Central Florida
Orlando, FL 32816-1364, U.S.A.
E-mail: piotrm@mail.ucf.edu
mtaylor@pegasus.cc.ucf.edu

Received 17.11.2008
and in final form 13.1.2009 (1951)