ACTUARIAL RESEARCH CLEARING HOUSE
1994 VOL. 2

ESTIMATION OF A MULTIVARIATE COPULA

Jacques F. Carriere

The University of Manitoba
Dept. of Actuarial and Management Sciences
Winnipeg, Manitoba, Canada R3T 2N2

Key words and phrases: multivariate copula, kernel distribution estimators, measures of association.

ABSTRACT

Let $C(\boldsymbol{u})$ be the multivariate copula of a distribution function $H(\boldsymbol{x}) = C(\boldsymbol{F}(\boldsymbol{x}))$ where $\boldsymbol{F}(\boldsymbol{x}) = (F_1(x_1),\ldots, F_p(x_p))^T$ are continuous marginal distributions. Given a random sample $\boldsymbol{X}_i$ for $i = 1,\ldots,n$ we will construct an estimate $\tilde{C}(\boldsymbol{u})$ based on kernel distribution estimators of $H(\boldsymbol{x})$ and $\boldsymbol{F}(\boldsymbol{x})$ and we will show that for all $\boldsymbol{u} \in R^p$, $\tilde{C}_n(\boldsymbol{u}) \to C(\boldsymbol{u})$ a.e. as $n \to \infty$.

1. INTRODUCTION

Let $p \geq 1$ be an integer and let $X = (X_1,\ldots,X_p)^T$ be a random vector that maps a probability space $(\Omega, \mathfrak{F}, P)$ into $(R^p, \mathfrak{B}^p)$ where $\mathfrak{B}^p$ are the Borel sets of the $p-$dimensional Euclidean space R^p. The distribution of X evaluated at $\boldsymbol{x} = (x_1,\ldots,x_p)^T \in R^p$ will be denoted as $H(\boldsymbol{x}) = P(X \leq \boldsymbol{x})$, where $X \leq \boldsymbol{x}$ if and only if $X_k \leq x_k$ $\forall$ $k=1,\ldots,p$. The marginals of $H(\boldsymbol{x})$ will be denoted as $F_k(x_k)= P(X_k \leq x_k)$ for $k=1,\ldots,p$. We suppose that $H(\boldsymbol{x})$ is continuous $\forall$ $\boldsymbol{x} \in R^p$. We start the discussion with a lemma about continuous distribution functions that is useful in the ensuing discussion.

Lemma 1.1. The following three conditions are equivalent for any $p \geq 1$:

i) $H(\mathbf{x})$ is continuous $\forall\ \mathbf{x} \in R^p$,

ii) $H(\mathbf{x})$ is uniformly continuous on R^p,

iii) $F(\mathbf{x})$ is uniformly continuous on R^p.

Proof: It is obvious that ii) implies i). That iii) implies ii) follows from the inequality $|H(\mathbf{x}) - H(\mathbf{y})| \leq \sum_{k=1}^{p} |F_k(x_k) - F_k(y_k)|$. This well known inequality may be found in Schweizer and Sklar (1983, p. 82). It is well known that $F(\mathbf{x})$ is uniformly continuous on R^p whenever $F(\mathbf{x})$ is continuous $\forall \mathbf{x} \in R^p$. Therefore, it is sufficient to show that i) implies that $F(\mathbf{x})$ is continuous $\forall \mathbf{x} \in R^p$. Let $\epsilon > 0$ and x, $x^* \in R$. Let $\mathbf{v}_k(x) = (x^*, \ldots, x^*, x, x^*, \ldots, x^*)^T$ be a vector with the $k-$th coordinate equal to x and all other coordinates equal to x^*. There exists $\delta > 0$ such that if $\mathbf{h} \in R^p$ and $\|\mathbf{h}\| = \underset{1 \leq k \leq p}{\mathrm{Max}} |h_k| \leq \delta$ then $|H(\mathbf{v}_k(x) + \mathbf{h}) - H(\mathbf{v}_k(x) - \mathbf{h})| < \epsilon/3$. Also there exists $x^* \in R$ such that $|F_k(x+\delta) - H(\mathbf{v}_k(x+\delta))| < \epsilon/3$ and $|F_k(x-\delta) - H(\mathbf{v}_k(x-\delta))| < \epsilon/3$. Therefore, if $|h| \leq \delta$ then $|F_k(x+h) - F_k(x-h)| \leq |F_k(x+\delta) - H(\mathbf{v}_k(x+\delta))| + |F_k(x-\delta) - H(\mathbf{v}_k(x-\delta))| + |H(\mathbf{v}_k(x+\delta)) - H(\mathbf{v}_k(x-\delta))| \leq \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$. So $F_k(x)$ is continuous $\forall\ x \in R$ and $\forall\ k=1, \ldots, p$. ■

Let $\mathbf{u} \in R^p$. Define $C(\mathbf{u}) = P(F(\mathbf{X}) \leq \mathbf{u})$. Then $C(\mathbf{u})$ is a distribution function with uniform marginals. Lemma 1.1 states that $C(\mathbf{u})$ is uniformly continuous on R^p because the marginals are continuous. Schweizer and Sklar (1983) call the function $C(\mathbf{u})$ a $p-$dimensional copula. An example of a copula is $C(\mathbf{u}) = \prod_{k=1}^{p} u_k$ and another is $C(\mathbf{u}) = \mathrm{Min}(u_1, \ldots, u_p)$ where $\mathbf{u} \in [0, 1]^p$. Examples of $2-$dimensional copulas may be found in Barnett (1980). Note that a copula relates a multivariate distribution function to its marginals. That is

$$H(\mathbf{x}) = C(F(\mathbf{x})). \tag{1.1}$$

This identity is true because $F(\mathbf{x})$ is uniformly continuous on R^p and so $P(\mathbf{X} \leq \mathbf{x}) = P(F(\mathbf{X}) \leq F(\mathbf{x})) = C(F(\mathbf{x}))$. For $u \in [0, 1]$, define $F_k^{-1}(u) = \inf\{x \in [-\infty, \infty]: F(x) \geq u\}$. For $\mathbf{u} \in [0, 1]^p$, define $F^{-1}(\mathbf{u}) = (F_1^{-1}(u_1), \ldots, F_p^{-1}(u_p))^T$. Then another useful identity is

$$C(\mathbf{u}) = H(F^{-1}(\mathbf{u})). \tag{1.2}$$

This identity is true because $F \circ F^{-1}(u) = u$. We now present some results about kernel distribution estimators of distribution functions.

2. KERNEL DISTRIBUTION ESTIMATORS

Let $X_i = (X_{1i},\ldots,X_{pi})^T$ for $i=1,2,\ldots$ be a sequence of independent and identically distributed random vectors each with a distribution equal to $H(x)$. With a finite sample $X_1,\ldots,X_n$, we can estimate $H(x)$ with the empirical distribution function

$$\hat{H}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathrm{I}(X_i \leq x). \tag{2.1}$$

For any $x \in R^p$, $\hat{H}_n(x)$ converges to $H(x)$ almost everywhere (a.e.). This implies that $\hat{H}_n \Rightarrow H$ a.e. as $n \to \infty$ where the notation $\Rightarrow$ means that the sequence of distribution functions converges weakly. Let $\delta_0(x) = \mathrm{I}(x \geq \mathbf{0})$ denote the distribution function of a measure that assigns unit mass at $\mathbf{0}$. Let $K_n(x)$ for $n=1,2,\ldots$ be a sequence of distribution functions such that $K_n \Rightarrow \delta_0$ as $n \to \infty$. This definition of a kernel sequence is a generalization of one presented in Rao (1983). A trivial example of a kernel sequence is one where $K_n(x) = \delta_0(x)\ \forall\ n=1,2,\ldots$ A kernel distribution estimator of $H(x)$ is

$$\tilde{H}_n(x) = \int_{R^p} K_n(x-y)\, d\hat{H}_n(y). \tag{2.2}$$

That is, $\tilde{H}_n(x)$ is the convolution of $\hat{H}_n(x)$ and $K_n(x)$. The following is a generalization of a result found in Rao (1983). This result will be useful in proving the main result given in Theorem 3.1. Note that $H(x)$ is not necessarily continuous in the following lemma.

Lemma 2.1. $\tilde{H}_n \Rightarrow H$ a.e. as $n \to \infty$.

Proof: Let $g(x)$ be bounded and continuous $\forall x \in R^p$. Then $\int_{R^p} g(x) d\tilde{H}_n(x) = \int_{R^p}\int_{R^p} g(x+y) dK_n(y) d\hat{H}_n(x)$. Let $g_n(x) = \int_{R^p} g(x+y) dK_n(y)$. Then by the definition of a kernel sequence $g_n(x) \to g(x)$ as $n \to \infty$. We know that if $A \in \mathfrak{B}^p$ then by the strong law of large numbers $\hat{H}_n(A) \to H(A)$ a.e. as $n \to \infty$. So by a generalized Lebesgue convergence theorem (Royden, 1968, p. 232) this implies that $\int_{R^p} g_n(x) d\hat{H}_n(x) \to \int_{R^p} g(x) dH(x)$ a.e. as $n \to \infty$. ∎

We now present a corollary that is useful for proving Theorem 3.1. Let $\tilde{F}_{kn}(x)$ for $k=1,\ldots,p$ be the marginal distributions of the kernel distribution estimator $\tilde{H}_n(\mathbf{x})$. Note that $\tilde{F}_{kn}(x)$ is itself a kernel distribution estimator of $F_k(x)$. Let $\tilde{\mathbf{F}}_n(\mathbf{x}) = (\tilde{F}_{1n}(x_1),\ldots, \tilde{F}_{pn}(x_p))^T$.

Corollary 2.2. Suppose $H(\mathbf{x})$ is uniformly continuous on R^p. Then $\sup_{\mathbf{x}\in R^p} |\tilde{H}_n(\mathbf{x}) - H(\mathbf{x})| \to 0$ a.e. and $\sup_{\mathbf{x}\in R^p} \|\tilde{\mathbf{F}}_n(\mathbf{x}) - \mathbf{F}(\mathbf{x})\| \to 0$ a.e. as $n \to \infty$.

Proof: Using a generalization of Polya's Theorem (Rao, 1962), we know that if $H(\mathbf{x})$ is uniformly continuous on R^p and $\tilde{H}_n \Rightarrow H$ a.e. as $n \to \infty$ then $\sup_{\mathbf{x}\in R^p} |\tilde{H}_n(\mathbf{x}) - H(\mathbf{x})| \to 0$ a.e. as $n \to \infty$. This is true for any $p \geq 1$. So for each $k=1,\ldots,p$ $\sup_{x\in R} |\tilde{F}_{kn}(x) - F_k(x)| \to 0$ a.e. as $n \to \infty$. ■

3. AN ESTIMATOR OF A COPULA

We now show how to estimate a $p-$dimensional copula with kernel distribution estimators. For $u \in [0, 1]$ define $\tilde{F}_{kn}^{-1}(u) = \inf\{x \in [-\infty, \infty]: \tilde{F}_{kn}(x) \geq u\}$. For $\mathbf{u} \in [0, 1]^p$ define $\tilde{\mathbf{F}}_n^{-1}(\mathbf{u}) = (\tilde{F}_{1n}^{-1}(u_1),\ldots, \tilde{F}_{pn}^{-1}(u_p))^T$. Using the identity $C(\mathbf{u}) = H(\mathbf{F}^{-1}(\mathbf{u}))$ presented in equation (1.2), we define our estimator as

$$\tilde{C}_n(\mathbf{u}) = \tilde{H}_n\left(\tilde{\mathbf{F}}_n^{-1}(\mathbf{u})\right). \tag{3.1}$$

We now show that under certain conditions on $H(\mathbf{x})$ and the kernel $K_n(\mathbf{x})$, the estimator $\tilde{C}_n(\mathbf{u})$ converges weakly. The major theorem of this paper now follows.

Theorem 3.1. Suppose $H(\mathbf{x})$ and $K_n(\mathbf{x})$ are continuous $\forall \mathbf{x} \in R^p$. Then $\tilde{C}_n(\mathbf{u})$ is a copula and $\tilde{C}_n \Rightarrow C$ a.e. as $n \to \infty$.

Proof: If $K_n(\mathbf{x})$ is continuous $\forall \mathbf{x} \in R^p$ then $\tilde{H}_n(\mathbf{x})$ is continuous $\forall \mathbf{x} \in R^p$. Therefore, by Lemma 1.1 $\tilde{\mathbf{F}}_n(\mathbf{x})$ is continuous $\forall \mathbf{x} \in R^p$. This means that the marginals of $\tilde{C}_n(\mathbf{u})$ are uniformly distributed and so $\tilde{C}_n(\mathbf{u})$ is a copula. Let $g : [0, 1]^p \to R$ be continuous. Then $g(\mathbf{u})$ is bounded and uniformly continuous on $[0, 1]^p$. We need to show that $\int_{[0,1]^p} g(\mathbf{u})d\tilde{C}_n(\mathbf{u}) \to \int_{[0,1]^p} g(\mathbf{u})dC(\mathbf{u})$ a.e. as $n \to \infty$. This is equivalent to showing that $\int_{R^p} g(\tilde{\mathbf{F}}_n(\mathbf{x}))d\tilde{H}_n(\mathbf{x}) \to \int_{R^p} g(\mathbf{F}(\mathbf{x}))dH(\mathbf{x})$ a.e. as $n \to \infty$. From Lemma

2.1 we know that $\int_{R^p} g(F(x))d\tilde{H}_n(x) \to \int_{R^p} g(F(x))dH(x)$ a.e. as $n \to \infty$ because $g(F(x))$ is continuous and bounded on R^p. All we need to show is that $\int_{R^p} |g(F(x)) - g(\tilde{F}_n(x))| \, d\tilde{H}_n(x) \to 0$ a.e. as $n \to \infty$. This will happen if we can show that $\sup_{x \in R^p} |g(F(x)) - g(\tilde{F}_n(x))| \to 0$ a.e. as $n \to \infty$. By the uniform continuity of $g(u)$ there exists $\delta > 0$ such that if $\|u_1 - u_2\| < \delta$ then $|g(u_1) - g(u_2)| < \epsilon$. By Corollary 2.2 there exists N such that $\forall n \geq N$ $\|\tilde{F}_n(x) - F(x)\| < \delta$ $\forall$ $x \in R^p$. So $\forall x \in R^p$ and $\forall n \geq N$ $|g(F(x)) - g(\tilde{F}_n(x))| < \epsilon$. ■

An immediate application of Theorem 3.1 occurs when the marginals $F(x)$ are known. Define

$$\tilde{\tilde{H}}_n(x) = \tilde{C}_n(F(x)). \tag{3.2}$$

Then the marginals of $\tilde{\tilde{H}}_n(x)$ are equal to $F(x)$ and $\tilde{\tilde{H}}_n \Rightarrow H$ a.e. as $n \to \infty$.

4. ESTIMATORS FOR CORRELATION COEFFICIENTS

We now show how to apply our results to the estimation of correlation coefficients. Suppose $H(x)$ is a 2 – dimensional copula. Consider Kendall's correlation coefficient equal to

$$\tau(H) = 4 \int_{R^2} H(x)dH(x) - 1. \tag{4.1}$$

Corollary 4.1. Suppose $H(x)$ is continuous $\forall x \in R^2$. Then $\tau(\tilde{H}_n) \to \tau(H)$ a.e. as $n \to \infty$.

Proof: From Lemma 2.1 we know that $\int_{R^2} H(x)d\tilde{H}_n(x) \to \int_{R^2} H(x)dH(x)$ a.e. as $n \to \infty$ because $H(x)$ is bounded and continuous $\forall x \in R^2$. Applying corollary 2.2, we find that $\int_{R^2} |\tilde{H}_n(x) - H(x)| \, d\tilde{H}_n(x) \leq \sup_{x \in R^2} |\tilde{H}_n(x) - H(x)| \to 0$ a.e. as $n \to \infty$. So $\int_{R^2} \tilde{H}_n(x)d\tilde{H}_n(x) \to \int_{R^2} H(x)dH(x)$ a.e. as $n \to \infty$. ■

Now consider Spearman's correlation coefficient equal to

$$\rho(C) = 12\int_{[0,1]^2} uv\ dC(u,\ v) - 3. \tag{4.2}$$

Corollary 4.2. Suppose $H(x)$ and $K_n(x)$ are continuous $\forall x \in R^p$. Then $\rho(\tilde{C}_n) \rightarrow \rho(C)$ a.e. as $n \rightarrow \infty$.

Proof. The function $g(u,v) = uv$ is continuous on $[0,\ 1]^2$. So by Theorem 3.1, $\int_{[0,1]^2} uv\ d\tilde{C}_n(u,\ v) \rightarrow \int_{[0,1]^2} uv\ dC(u,\ v)$ a.e. as $n \rightarrow \infty$. Therefore, $\rho(\tilde{C}_n) \rightarrow \rho(C)$ a.e. as $n \rightarrow \infty$. ∎

BIBLIOGRAPHY

Barnett, V. (1980). "Some Bivariate Uniform Distributions." *Commun. Statist. – Theor. Method A*, **9**, 453 – 461.

Rao, B.L.S.P. (1983). *Non-Parametric Functional Estimation*. New York: Academic Press.

Rao, R. (1962). "Relations between weak and uniform convergence of measurement with applications." *Ann. Math. Statist.*, **33**, 659 – 681.

Royden, H. L. (1968). *Real Analysis*. New York: MacMillan.

Schweizer, B. and Sklar, A. (1983). *Probabilistic Metric Spaces*. New York: North Holland.