EDITORIAL

This special issue is devoted to the applications of the theory of copulas in probability and statistics to various important problems in econometrics.

From a mathematical viewpoint, copula (a Latin word for "link") should be a very old and simple concept in probability theory. Indeed, a copula is simply a multivariate distribution function (of a random vector) whose marginals are uniform distributions on the unit interval $[0, 1]$. However, this concept was emphasized only in Abe Sklar's thesis (1959) in response to an earlier question by Maurice Frechet concerning joint distributions with given marginals.

So why such a simple concept turns out to be so essential in probability and statistics, leading to more realistic models in almost all empirical sciences, including economics, as we witness today?

In real analysis, or more specifically, in measure theory, we are almost exclusively concerned with a product measure built from measures on factor spaces. Lebesgue measure on $\mathbb{R}^2$ is simply the product measure of Lebesgue measures on $\mathbb{R}$. We might not have compelling reasons to seek other kinds of measures on $\mathbb{R}^2$, built also from Lebesgue measure on $\mathbb{R}$. More generally, given two measure spaces $(\Omega_1, \mathcal{A}_1, \mu_1)$, $(\Omega_2, \mathcal{A}_2, \mu_2)$, are we concerned with all possible measures on $\mathcal{A}_1 \otimes \mathcal{A}_2$ whose projections on $\mathcal{A}_1, \mathcal{A}_2$ are μ_1, μ_2, respectively?

As we will see, the answer is definitely yes in the special case of probability measures, for various reasons.

As far as random vectors (as a specific type of random elements) are concerned, their probability laws on $\mathbb{R}^d$ are characterized by Lebesgue-Stieltjes Theorem: there is an one-to-one and onto (a bijection) map between probability measures on $\mathcal{B}(\mathbb{R}^d)$ and (multivariate) distribution functions on $\mathbb{R}^d$. This characterization theorem is very important for applications since it provides a simpler way to obtain statistical models. For example, a bivariate (real-valued) random vector (X, Y) with both X and Y being defined on the same probability space $(\Omega, \mathcal{A}, P)$, is characterized by a joint distribution function $H(., .) : \mathbb{R}^2 \to [0, 1]$ where $H(x, y) = P(X \leq x, Y \leq y)$, in the sense that H determines the probability measure (law) of (X, Y) on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$. Statistically speaking, the joint distribution H contains all information about (X, Y). For example, the marginals are derived as $F(x) = P(X \leq x) = H(x, \infty)$, $G(y) = P(Y \leq y) = H(\infty, y)$. Their associated probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, $dF(x)$ and $dG(y)$ are marginal measures of $dH(x, y)$. However, unless X and Y are independent, the "joint" probability measure $dH(x, y)$ might not be the product measure $dF(x) \otimes dG(y)$, or equivalently the relation between $H(.,.)$ and its marginals is $H(x, y) = F(x)G(y)$ for all $x, y \in \mathbb{R}$.

It is the notion of "stochastic dependence" which is central in using statistics to discover knowledge, at the service of science. But as $H(x, y)$ contains all information about the random evolution of (X, Y), it should provide information about the dependence between X and Y. Well, it suffices to calculate the conditional distribution, say, of Y given X, namely $F_{Y|X=x}(y) = P(Y \leq x|X = x)$, and compare with the unconditional distribution $G(y)$. The result will tell us whether X and Y are independent or not. In case where the result reveals that they are not independent, can we learn more about "what kind of dependence structure they possess?". As we will see, such a question turns out to be of great interest in applications. Here, you might remember what mathematicians used to ask "If we are not doing linear mathematics, what kinds of nonlinear mathematics we are doing?". Since we do not have an answer to this, we just have two areas of mathematics: linear and nonlinear mathematics! Thanks to Sklar, we have an answer to "If random variables are not (mutually) independent, what kind of dependence they possess?". The answer is copulas! Each copula represents one type of dependence. But why this important question was not asked earlier than Maurice Frechet? and, in fact, even after Sklar got the answer in 1959, it did not seem to "revolutionize" statistics until the 1990s? The reasons are manyfold. First, based on "tradition", joint distributions are in general given, such as multivariate Gaussian distributions, rather than need to be built from given marginals. Secondly, we used to rely modeling on conditional models, i.e., building the joint distributions from marginals and conditional distributions. And thirdly, as far as dependence is concerned, by tradition again, statisticians focus right away on quantifying dependence rather than looking for the qualitative concept of dependence. Correlation analysis dominated statistical analyses since the beginning. Even it is clear that correlation is not causation, statistical analyses focus only on quantifying relationship between variables. This resulted in looking mainly at one type of dependence, namely linear dependence (quantified by the Pearson coefficient correlation). Two exceptions are Kandell tau and Spearman rho, quantifying two other types of dependence structures, namely comonotonicity and counter-comonotonicity, respectively. But more importantly, without knowing how to capture (represent) a type of suspected dependence, statisticians just go ahead to sssume independence, such as modeling relationships between random shocks between various statistical models.

The appearance of copulas has changed all the above, leading to much more realistic statistical modeling in applications, in all fields where statistics is the main tool for investigation. The point is this. We need innovation!

The "hidden" thing seems to be this. How to relate a joint distribution H to is marginals F and G? In other words, can we pose the inverse problem: Given F and G, what can be said about joint distributions which admit F and G as marginals? Or more specifically, can we determine all possible such joint distributions?

From a given joint distribution $H(.,.)$ of (X,Y), it is not obvious how to "find" the type of dependence between X and Y. Now, since F and G are nondecreasing,

$$H(x,y) = P(X \leq x, Y \leq y) = P(F(X) \leq F(x), G(Y) \leq G(y))$$

and if F and G are continuous, then the random variable $U = F(X)$, $V = G(Y)$ are uniformly distributed on $[0,1]$, so that

$$H(x,y) = C(F(x), G(y))$$

where C denotes the joint distribution of (U,V), i.e., C is a joint distribution with uniform marginals on $[0,1]$. C is called a (bivariate) copula. Of course, the result is true for any dimensions n, but for simplicity, we just consider the case $n = 2$. As a such joint distribution, when restricted to its support $[0,1]^2$, a copula is defined as a function $C : [0,1]^2 \to [0,1]$ satisfying the following axioms:

(i) $C(u,0) = C(0,v) = 0$ for any $u, v \in [0,1]$

(ii) $C(u,1) = u$, $C(1,v) = v$ for any $u, v \in [0,1]$

(iii) For any $u \leq u', v \leq v'$,

$$C(u',v') - C(u',v) - C(u,v') + C(u,v) \geq 0$$

The upshot is Sklar's theorem which says that the above is in fact general. If H is a bivariate distribution with marginals F and G, then there exists a copula C such that, for all $x, y \in \bar{\mathbb{R}}$, $H(x,y) = C(F(x), G(y))$. If F and G are both continuous, then C is unique, otherwise, it is determined on the range $\mathcal{R}(F) \times \mathcal{R}(G)$. Conversely, if C is a copula, and F and G are marginal distributions, then $H(x, y) = C(F(x), G(y))$ is a joint distribution having F and G as its marginals.

As a result, we can extract the copula from the joint distribution: $C(u, v) = H(F^{-1}(u), G^{-1}(v))$, where $F^{-1} : [0, 1] \to \mathbb{R}$ is the quantile function $F^{-1}(u) = \inf\{x \in \mathbb{R} : F(x) \geq u\}$.

The implications of Sklar's results are of the following types. In applications, if we have only two marginals F and G (estimated from data), then to model the process generating the joint observations, we just need to "look for" an appropriate copula. If we have a joint distribution, then extracting its copula will reveal the type of dependence. Of course, the next task will be "how to quantify a copula?", i.e., assessing the strength of the dependence expressed by a copula.

Copula theory has become a norm in econometrics. This special issue consists of a sample of important applications in econometrics using copula methodology.

Sompong Dhompongsa
Hung T. Nguyen
Chiang Mai, August 2014