SFB 823

Almost opposite regression dependence in bivariate distributions

Karl Friedrich Siburg, Pavel A. Stoimenov

Nr. 27/2014

Discussion Paper

Almost opposite regression dependence in bivariate distributions

Karl Friedrich Siburg
Fakultät für Mathematik
Technische Universität Dortmund
Vogelpothsweg 87
44227 Dortmund
Germany
Tel: +49-231-755 7211 (Fax: 5219)
Email: karl.f.siburg@math.tu-dortmund.de

Pavel A. Stoimenov
International Energy Company
Vienna
Austria
Email: pavel.stoimenov@gmail.com

August 5, 2014

Abstract

Let X,Y be two continuous random variables. Investigating the regression dependence of Y on X, respectively, of X on Y, we show that the two of them can have almost opposite behavior. Indeed, given any $\varepsilon > 0$, we construct a bivariate random vector (X,Y) such that the respective regression dependence measures $r_{2|1}(X,Y), r_{1|2}(X,Y) \in [0,1]$ introduced in Dette et al. (2013) satisfy $r_{2|1}(X,Y) = 1$ as well as $r_{1|2}(X,Y) < \varepsilon$.

1 Introduction and results

Recently, Dette et al. (2013) introduced an approach to the problem of ordering and measuring regression dependence in the bivariate case. Let (X,Y) be a bivariate random vector. Since regression dependence is a directional relationship, it is first necessary to specify the direction of interest. Without loss of generality, consider the dependence of Y on X. The fundamental idea behind regression is predictability—the more predictable Y is from X, the more regression dependent they are. It is straightforward to single out the two extreme cases: independence and almost sure functional dependence, when there exists a Borel measurable function g such that $Y = g(X)$ with probability one (Lancaster, 1963). In the former case, X provides no information about Y, whereas in the latter case there is perfect predictability of Y from X.

Apart from the two extreme cases, however, there exists a variety of intermediate ones with a certain degree of regression dependence. To be able to measure the strength of dependence of Y on X, Dette et al. (2013) introduced the concept of an order of regression dependence. Only such an order allows one to deal with questions like whether one random variable Y can be better regressed onto X than another random variable Y' (namely when (X,Y) is more regression dependent than (X,Y')). Note that the concept of regression dependence is quite different from the well known concept of dependence, as measured by a variety of measures of dependence or association. Indeed, the general notion of dependence is not a directional concept, i.e., it cannot describe how strongly Y depends on X.

In addition to an order of regression dependence, Dette et al. (2013) constructed a nonparametric measure of regression dependence, $r_{2|1}(X,Y) \in [0,1]$, which is monotone in this order. Moreover, the measure takes on its extreme values precisely at independence and almost sure functional dependence, respectively, i.e., we have

(i) $r_{2|1}(X,Y) = 1$ if and only if Y is a.s. a Borel function of X.

(ii) $r_{2|1}(X,Y) = 0$ if and only if X and Y are independent.

We point out that it is important to have equivalences in both of the properties (i) and (ii), because only then the value $r_{2|1}(X,Y)$ can serve as a genuine measure of how much Y is dependent on X. Indeed, if we only had $r_{2|1}(X,Y) = 0$ if (but not only if) X and Y are independent, then an assertion like $r_{2|1}(X,Y) < \varepsilon$ would not imply that Y is 'almost independent' from X.

Analogously, of course, one can exchange the roles of X on Y and define a measure $r_{1|2}(X,Y) = r_{2|1}(Y,X)$ measuring the degree of dependence of X on Y.

The following is the main result of the present paper.

Theorem 1. *For any given $\varepsilon > 0$, there is a random vector (X,Y) such that the following assertions hold:*

1. $r_{2|1}(X,Y) = 1$, *i.e., Y is a.s. a Borel function of X.*
2. $r_{1|2}(X,Y) < \varepsilon$.

The paper is organized as follows. In Section 2 we give a quick review of the construction in Dette et al. (2013) of the nonparametric measure $r_{2|1}$ of regression dependence. Section 3 then contains the proof of Theorem 1. Section 4 relates this result to other problems in the literature.

Acknowledgements: Research on this paper started when PAS was supported by the DFG Sonderforschungsbereich 823 "'Statistical modelling of nonlinear dynamic processes"'.

2 Preliminaries

In this section we recall the basic notion of copula and the definition of the nonparametric measure of regression dependence introduced in Dette et al. (2013). A (two-dimensional) copula is a function $C : I^2 \to I$ with $I := [0,1]$, satisfying the following conditions:

1. $C(x,0) = C(0,y) = 0$ for all $x,y \in I$
2. $C(x,1) = x$ and $C(1,y) = y$ for all $x,y \in I$
3. C is 2-increasing, i.e., $C(x_2,y_2) - C(x_2,y_1) - C(x_1,y_2) + C(x_1,y_1) \geq 0$ for all rectangles $[x_1,x_2] \times [y_1,y_2] \subset I^2$.

These conditions imply further key properties. A copula is Lipschitz continuous and increasing in each argument; therefore, its partial derivatives exist a.e. on I^2. We refer the reader to Nelsen (2006) for more information about copulas.

Given two continuous random variables X and Y with corresponding copula C, the measure of regression dependence $r_{2|1}(X,Y)$ introduced in Dette et al. (2013) is defined by

$$r_{2|1}(X,Y) = 6\|\partial_1 C\|_2^2 - 2 = 6\int_{I^2} |\partial_1 C(x,y)|^2 \mathrm{d}(x,y) - 2 \tag{1}$$

where ∂_1 denotes the partial derivative with respect to the first variable and $\|\cdot\|_2$ is the L^2-norm on I^2. The quantity $r_{2|1}$ measures the degree of dependence of Y on X. It is a measure of regression dependence with respect to two natural regression dependence orders, also introduced in Dette et al. (2013).

Analogously, one can define a measure

$$r_{1|2}(X,Y) = 6\|\partial_2 C\|_2^2 - 2 = r_{2|1}(Y,X)$$

such that this quantity measures the degree of dependence of X on Y.

3 Proof of Theorem 1

In this section, we will construct a sequence (X_n, Y_n) of bivariate random vectors such that

$$r_{2|1}(X_n, Y_n) = 1 \text{ for all } n, \tag{2}$$

$$\lim_{n\to\infty} r_{1|2}(X_n, Y_n) = 0. \tag{3}$$

This proves Theorem 1. In fact, we will construct a sequence of copulas C_n rather than the random variables themselves. This is sufficient because the measures $r_{2|1}$ and $r_{1|2}$ depend only on the corresponding copula. For the construction of these copulas, we use the so-called gluing method developed in Siburg and Stoimenov (2008); alternatively, one could also use the orthogonal grid construction described in De Baets and De Meyer (2007). For the convenience of the reader, we quickly recall the details of the gluing method.

Given two copulas C_1, C_2 and a parameter $\theta \in (0,1)$, we define the function

$$(C_1 \circledast_{x=\theta} C_2)(x,y) = \begin{cases} \theta C_1\left(\frac{x}{\theta}, y\right) & \text{if } 0 \le x \le \theta \\ (1-\theta) C_2\left(\frac{x-\theta}{1-\theta}, y\right) + \theta y & \text{if } \theta \le x \le 1 \end{cases} \tag{4}$$

Thus, $C_1 \circledast_{x=\theta} C_2$ corresponds to gluing the two copulas C_1 and C_2: it equals C_1, rescaled and fit into the rectangle $[0,\theta] \times I$, and equals $C_2 + \theta y$, rescaled and fit into $[\theta, 1] \times I$. It is shown in Siburg and Stoimenov (2008) that the gluing process yields a copula again, i.e., $C_1 \circledast_{x=\theta} C_2$ is a copula for any parameter θ. For later purposes, we need also the gradient of the resulting copula which is given by

$$\begin{aligned} &\nabla(C_1 \circledast_{x=\theta} C_2)(x,y) \\ &\quad = \begin{cases} \left(\partial_1 C_1\left(\frac{x}{\theta}, y\right), \theta \partial_2 C_1\left(\frac{x}{\theta}, y\right)\right) & \text{if } 0 \le x \le \theta \\ \left(\partial_1 C_2\left(\frac{x-\theta}{1-\theta}, y\right), (1-\theta)\partial_2 C_2\left(\frac{x-\theta}{1-\theta}, y\right) + \theta\right) & \text{if } \theta \le x \le 1 \end{cases} \end{aligned} \tag{5}$$

provided the partial derivatives on the right exist.

Let us first illustrate the glueing construction with a fundamental example. Recall that a copula C is called singular if its density $\partial^2 C/\partial x \partial y$ vanishes almost everywhere in I^2. Moreover, the support of a copula C is defined as the complement of the union of all (relatively) open subsets of I^2 whose measure, induced by C, is zero. We refer to Nelsen (2006) for more details.

Example 1. Let $\theta \in (0,1)$, and suppose that the probability θ is uniformly distributed along the line segment joining $(0,0)$ and $(\theta,1)$, and the probability $1-\theta$ is uniformly distributed along the segment between $(\theta,1)$ and

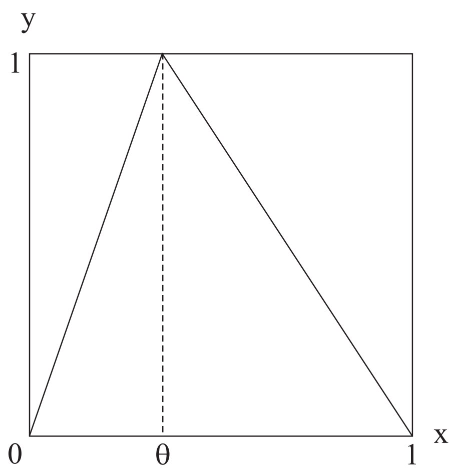

Figure 1: The support of the singular copula C_θ in Example 1

$(1,0)$. Consider the resulting singular copula C_θ whose support consists of these two line segments; see Figure 1. It follows (see (Nelsen, 2006, Ex. 3.3)) that

$$C_\theta(x,y) = \begin{cases} x & \text{if } x \leq \theta y \\ \theta y & \text{if } \theta y < x < 1-(1-\theta)y \\ x+y-1 & \text{if } 1-(1-\theta)y \leq x. \end{cases}$$

Note that C_θ can be written as the gluing

$$C_\theta = C^+ \circledast_{x=\theta} C^-$$

where $C^+(x,y) = \min(x,y)$ and $C^-(x,y) = \max(x+y-1,0)$ are the upper and lower Fréchet-Hoeffding bound, respectively.

Since the support of C_θ is a graph over the x-axis, this copula links random variables X and Y where Y is completely dependent on X. This follows from Dette et al. (2013, Prop. 1) and the fact that a function is Borel measurable if and only if its graph is Borel measurable and has probability one (Buckley, 1974). On the other hand, X is not completely dependent on Y because the support of C_θ is not a graph over the y-axis.

This example will serve as a fundamental building block for our final construction of copulas C_n satisfying (2) and (3). To do so, we start with the copula $C^+ \circledast_{x=\theta} C^-$ from Example 1 where, in order to simplify calculations, we set $\theta = 1/2$. Then we define C_n inductively by

$$C_1 = C^+ \circledast_{x=1/2} C^-$$
$$C_{n+1} = C_n \circledast_{x=1/2} C_n$$

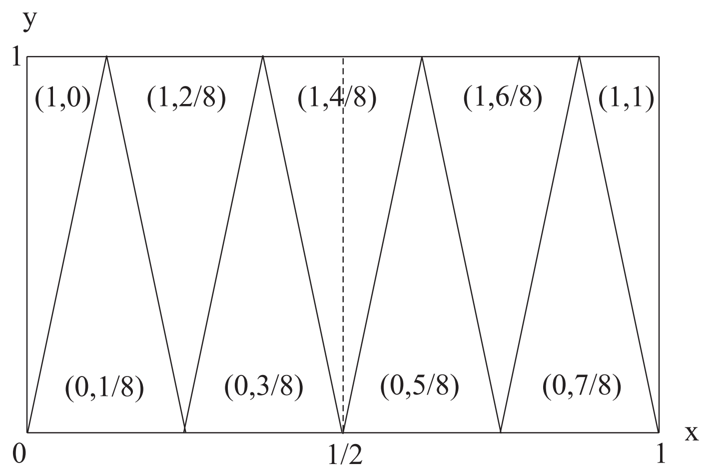

Figure 2: The gradient of the copula C_3 in the proof of Theorem 1

for $n \geq 1$. We claim that

$$\int_{I^2} |\partial_1 C_n(x,y)|^2 \mathrm{d}(x,y) = \frac{1}{2} \tag{6}$$

for all $n \geq 1$, as well as

$$\int_{I^2} |\partial_2 C_n(x,y)|^2 \mathrm{d}(x,y) \to \frac{1}{3} \tag{7}$$

as $n \to \infty$. These relations imply that

$$r_{2|1}(X,Y) = 6 \int_{I^2} |\partial_1 C_n(x,y)|^2 \mathrm{d}(x,y) - 2 = 1$$

for all n, as well as

$$r_{1|2}(X,Y) = 6 \int_{I^2} |\partial_2 C_n(x,y)|^2 \mathrm{d}(x,y) - 2 \to 0$$

as $n \to \infty$, which are precisely the assertions (2) and (3) that we wanted to prove.

For the proof of (6) and (7), we have to calculate the gradient ∇C_n. Using (5) and the fact that $1-\theta = \theta = 1/2$, we see that $\partial C_n/\partial x = 1$ in the upper and $\partial C_n/\partial x = 0$ in the lower triangles formed by the line segments of the support of C_n, and the second component $\partial C_n/\partial y$ takes the values $0, 1/2^n, 2/2^n, \ldots, (2^n-1)/2^n, 1$ respectively; see Figure 2 for the case $n = 3$.

Since the gradient of C_n is constant on each triangle, the integration reduces to multiplying the square of the respective constant with the area of the corresponding triangle. Thus, considering the first component of the gradient, we obtain

$$\int_{I^2} |\partial_1 C_n(x,y)|^2 \mathrm{d}(x,y) = \frac{1}{2}$$

for each $n \geq 1$, proving (6).

Now we deal with the second component of the gradient. Each of the triangles in Figure 2, except for the two triangles on the left and the right, has area $1/2^n$, and the value of the second component of the gradient is $i/2^n$ where i ranges from 1 to $2^n - 1$. Therefore, the integral for the second component amounts to

$$\int_{I^2} |\partial_2 C_n(x,y)|^2 \mathrm{d}(x,y) = \left[\sum_{i=1}^{2^n-1} \left(\frac{i}{2^n}\right)^2 \cdot \frac{1}{2^n}\right] + 1^2 \cdot \frac{1}{2^{n+1}}$$

where the last term stems from the triangle containing the vertex $(1,1)$ which is just half as big as the other ones. Using the formula

$$\sum_{i=1}^{k-1} i^2 = k^3/3 + \mathscr{O}(k^2)$$

we conclude that

$$\int_{I^2} |\partial_2 C_n(x,y)|^2 \mathrm{d}(x,y) = \frac{1}{2^{n+1}} + \left(\frac{1}{2^n}\right)^3 \cdot \sum_{i=1}^{2^n-1} i^2 = \frac{1}{3} + \mathscr{O}\left(\frac{1}{2^n}\right)$$

as $n \to \infty$, proving also our claim (7).

4 Further remarks

Given an abstract measure of regression dependence, say $\rho_{2|1}$, one could try to construct a measure of dependence by setting $\rho = (\rho_{2|1} + \rho_{1|2})/2$. Note that, a priori, it is not clear at all why this definition should yield a decent measure of dependence. However, if we consider the measures $r_{2|1}$ and $r_{1|2}$, then this idea does give a meaningful result. Indeed, the function

$$\begin{aligned}\omega(X,Y)^2 &= (r_{2|1}(X,Y) + r_{1|2}(X,Y))/2 \\ &= 3\int_{I^2} |\partial_1 C(x,y)|^2 + |\partial_2 C(x,y)|^2 \mathrm{d}(x,y) - 2\end{aligned}$$

is (the square of) the measure of mutual dependence ω introduced and studied in Siburg and Stoimenov (2010).

It was shown in (Siburg and Stoimenov, 2010, Thm. 13(vi)) that

$$\frac{\sqrt{2}}{2} \leq \omega(X,Y) \leq 1$$

whenever Y is a.s. a Borel function of X. However, it was not clear whether the lower bound was sharp. Our above example now shows that this is indeed the case. Namely, for the sequence of random variables (X_n, Y_n), respectively, their corresponding copulas C_n as in Section 3 we have

$$\lim_{n\to\infty} r_{2|1}(X_n, Y_n) + r_{1|2}(X_n, Y_n) = 1, \tag{8}$$

which implies that

$$\lim_{n\to\infty} \omega(X_n, Y_n) = \frac{\sqrt{2}}{2}.$$

References

Buckley, J. J. (1974). Graphs of measurable functions. *Proc. Amer. Math. Soc. 44*(1), 78–80.

De Baets, B. and H. De Meyer (2007). Orthogonal grid construction of copulas. *IEEE Trans. Fuzzy Systems 15*(6), 1053–1062.

Dette, H., K. F. Siburg, and P. A. Stoimenov (2013). A copula-based non-parametric measure of regression dependence. *Scand. J. Stat. 40*(1), 21–41.

Lancaster, H. O. (1963). Correlation and complete dependence of random variables. *Ann. Math. Statist. 34*, 1315–1321.

Nelsen, R. B. (2006). *An introduction to copulas* (second ed.). New York: Springer.

Siburg, K. F. and P. A. Stoimenov (2008). Gluing copulas. *Comm. Statist. Theory Methods 37*(19), 3124–3134.

Siburg, K. F. and P. A. Stoimenov (2010). A measure of mutual complete dependence. *Metrika 71*(2), 239–251.