Statistics and Probability Letters 169 (2021) 108960

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

On relationships between the Pearson and the distance correlation coefficients

Dominic Edelmann [a], Tamás F. Móri [b,*], Gábor J. Székely [c,b]

[a] Division of Biostatistics, German Cancer Research Center, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany
[b] Alfréd Rényi Institute of Mathematics, Reáltanoda u. 13–15, H-1053 Budapest, Hungary
[c] National Science Foundation, 2415 Eisenhower Avenue, Alexandria, VA 22314, USA

A R T I C L E I N F O

Article history:
Received 13 December 2019
Received in revised form 19 June 2020
Accepted 20 July 2020
Available online 5 October 2020

MSC:
62H20

Keywords:
Distance correlation
Pearson correlation
Distance covariance
Mixture distribution
Rademacher distribution

A B S T R A C T

In this paper we show that for any fixed Pearson correlation coefficient strictly between -1 and 1, the distance correlation coefficient can take any value in the open unit interval $(0, 1)$.

1. Introduction

Measuring dependence between random observations undoubtedly plays a central role in statistics. Since it is very difficult to fully understand and describe dependencies, one is often interested in condensing the strength of dependence into one single number. Such a number, which is typically defined on either of the intervals $[-1, 1]$ or $[0, 1]$ is called a *correlation coefficient*. The classical and arguably most popular correlation coefficient is the Pearson one, which, for random variables X and Y with finite and positive variances, is defined by

$$\mathrm{cor}(X, Y) := \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}},$$

where $\mathrm{cov}(X, Y)$ denotes the covariance of X and Y and for any random variable Z, $\mathrm{Var}(Z)$ denotes the variance of Z. By the Cauchy–Schwarz inequality, we have $\mathrm{cor}(X, Y) \in [-1, 1]$. It is also well-known that $|\mathrm{cor}(X, Y)| = 1$ if and only if $Y = aX + b$, almost surely, with real constants $a \neq 0$ and b and that the independence of X and Y implies $\mathrm{cor}(X, Y) = 0$. An important drawback of Pearson's correlation coefficient is that the converse implication is not true: It is straightforward to construct non-independent random variables X and Y, such that $\mathrm{cor}(X, Y) = 0$. A simple example is given any two random variables X and Y, where X is symmetrically distributed around zero and $Y = |X|$, cf. Section 2.

An alternative correlation coefficient, which does not suffer from this drawback is the *distance correlation coefficient* (Székely et al., 2007; Székely and Rizzo, 2009). Let X and Y be real valued random variables with finite second

* Corresponding author.
E-mail addresses: dominic.edelmann@dkfz-heidelberg.de (D. Edelmann), mori.tamas@renyi.hu (T.F. Móri), gszekely@nsf.gov (G.J. Székely).

https://doi.org/10.1016/j.spl.2020.108960
0167-7152/

moments. Then *the distance covariance* can be defined in the following form. Let (X, Y), (X', Y'), (X'', Y'') denote independent and identically distributed copies then the distance covariance is the square root of

$$\begin{aligned} \mathrm{dCov}^2(X, Y) := {} & \mathbb{E}(|X - X'|\,|Y - Y'|) + \mathbb{E}(|X - X'|)\mathbb{E}(|Y - Y'|) \\ & -2\mathbb{E}(|X - X'|\,|Y - Y''|). \end{aligned} \tag{1}$$

If the random variables X, Y have finite and positive variances, then the definition of their *distance correlation coefficient* is the following:

$$\mathrm{dCor}(X, Y) := \frac{\mathrm{dCov}(X, Y)}{\sqrt{\mathrm{dCov}(X, X)\mathrm{dCov}(Y, Y)}}. \tag{2}$$

The distance correlation coefficient is nonnegative with $\mathrm{dCor}(X, Y) \in [0, 1]$. Pearson's correlation, and also the distance correlation coefficient is invariant with respect to linear transformations, and $\mathrm{dCor}(X, Y) = 1$ if and only if $Y = aX + b$ almost surely, with real constants $a \neq 0$ and b (Székely et al., 2007). However, there are two crucial differences of the Pearson and the distance correlation coefficients. First, although this is not directly clear from Eq. (1), the definition of the distance correlation coefficient can be extended to variables with finite first moments only. A possible way to define the distance covariance under this less restrictive assumption is

$$\begin{aligned} \mathrm{dCov}^2(X, Y) := {} & \mathbb{E}\big(|X - X'| \cdot (|Y - Y'| - |Y - Y''| - |Y' - Y''|)\big) \\ & +\mathbb{E}(|X - X'|)\mathbb{E}(|Y - Y'|); \end{aligned} \tag{3}$$

the distance correlation coefficient is then defined *via* Eq. (2). The second important difference is that the lack of distance correlation defines independence; that is, $\mathrm{dCor}(X, Y) = 0$ if and only if X and Y are independent.

Since both the absolute value of Pearson's correlation coefficient and the distance correlation coefficient are used in applications to quantify strength of dependence, it is important to understand how large the differences between these two measures can possibly be. It is immediately clear that $\mathrm{dCor}(X, Y) = 0$ implies $\mathrm{cor}(X, Y) = 0$ and that $|\mathrm{cor}(X, Y)| = 1$ if and only if $\mathrm{dCor}(X, Y) = 1$. It is also straightforward to show that the distance correlation coefficient of random variables only having two possible values coincides with their absolute Pearson correlation coefficient. Moreover, several results for bivariate parametric distributions have been derived (Székely et al., 2007; Dueck et al., 2017). In the special case when (X, Y) are jointly distributed as bivariate normal, their distance correlation coefficient is a deterministic function of the Pearson correlation coefficient $\varrho = \mathrm{cor}(X, Y)$ (Székely et al., 2007, Theorem 7), namely,

$$\mathrm{dCor}^2(X, Y) = \frac{\varrho \arcsin \varrho + \sqrt{1 - \varrho^2} - \varrho \arcsin(\varrho/2) - \sqrt{4 - \varrho^2} + 1}{1 + \pi/3 - \sqrt{3}}.$$

Note that this is a strictly increasing, convex function of $|\varrho|$, $\mathrm{dCor}(X, Y) \leq |\varrho|$ with equality when $\varrho = 0$ or $\varrho = \pm 1$. However, none of this work considered establishing extremal cases featuring large differences between these two correlation coefficients. A natural question to ask is for example how large the distance correlation coefficient can be for two random variables which are *uncorrelated*; that is, have Pearson's correlation coefficient equal to 0.

In this article, we will consider two different strategies for constructing extremal examples. In Section 2, we will investigate the class of random vectors (X, Y), where X is symmetric around 0 and $Y = |X|$. While it is immediately clear that $\mathrm{cor}(X, Y) = 0$ and $\mathrm{dCor}(X, Y) > 0$, we will show that the distance correlation coefficient cannot be arbitrarily close to 1. Section 3 follows a different approach. The fact that dCor is defined for X and Y with finite first moments, while $\mathrm{cor}(X, Y)$ requires finite second moments leads us to the conjecture that the Pearson correlation coefficient is more sensitive to dependence in the tails than the distance correlation. This conjecture motivates the construction of a specific mixture distribution showing that, up to trivial exceptions, all possible values for cor and dCor can be simultaneously achieved.

2. Correlation and distance correlation—how much can they differ?

Since uncorrelatedness only means the lack of linear dependence, one can naturally ask how large the distance correlation coefficient of two uncorrelated random variables can be. To find an extremal construction it seems plausible to investigate the case where one of the random variables is symmetrically distributed around zero and the other one is its absolute value. Such a pair is obviously uncorrelated in the classical (Pearson) sense, and since they are equal with probability 1/2 and the negative of each other with probability 1/2, we can hope a large distance correlation coefficient.

In what follows we compute the distance correlation coefficient between a symmetrically distributed random variable and its absolute value.

Theorem 1. *The distance correlation coefficient of a symmetrically distributed random variable and its absolute value is less than $2^{-1/4}$ and this bound is sharp.*

Proof. Let X be a nonnegative random variable, $\mathbb{E}(X^2) < \infty$, and let ε be independent of X with $\mathbb{P}(\varepsilon = 1) = \mathbb{P}(\varepsilon = -1) = 1/2$ (a random sign). Again, let X, X', X'', ε, ε', and ε'' be independent, and the random variables denoted by the same letter no matter that without or with a prime or double prime are supposed to be identically distributed. Note that here εX is symmetrically distributed with absolute value X. Clearly,

$$\begin{aligned}\mathrm{dCov}^2(X, X) &= \mathbb{E}\big((X - X')^2\big) + \big(\mathbb{E}(|X - X'|)\big)^2 - 2\mathbb{E}(|X - X'|\,|X - X''|)\\ &= 2\mathbb{E}(X^2) - 2\big(\mathbb{E}X\big)^2 + \big(\mathbb{E}(|X - X'|)\big)^2 - 2\mathbb{E}(|X - X'|\,|X - X''|).\end{aligned}$$

Similarly,

$$\begin{aligned}&\mathrm{dCov}^2(\varepsilon X, \varepsilon X)\\ &\quad= \mathbb{E}\big((\varepsilon X - \varepsilon' X')^2\big) + \big(\mathbb{E}(|\varepsilon X - \varepsilon' X'|)\big)^2 - 2\mathbb{E}(|\varepsilon X - \varepsilon' X'|\,|\varepsilon X - \varepsilon'' X''|)\\ &\quad= \mathbb{E}\big((X - \varepsilon' X')^2\big) + \big(\mathbb{E}(|X - \varepsilon' X'|)\big)^2 - 2\mathbb{E}(|X - \varepsilon' X'|\,|X - \varepsilon'' X''|)\\ &\quad= \tfrac{1}{2}\,\mathbb{E}\big((X - X')^2\big) + \tfrac{1}{2}\,\mathbb{E}\big((X + X')^2\big) + \big(\tfrac{1}{2}\,\mathbb{E}(|X - X'|) + \mathbb{E}X\big)^2\\ &\qquad - \tfrac{1}{2}\,\mathbb{E}(|X - X'|\,|X - X''|) - \mathbb{E}(X|X - X'|) - \mathbb{E}X\,\mathbb{E}(|X - X'|)\\ &\qquad - \tfrac{1}{2}\,\mathbb{E}(X^2) - \tfrac{3}{2}\,\big(\mathbb{E}X\big)^2\\ &\quad= \tfrac{3}{2}\,\mathbb{E}(X^2) - \tfrac{1}{2}\,\big(\mathbb{E}X\big)^2 + \tfrac{1}{4}\,\big(\mathbb{E}(|X - X'|)\big)^2\\ &\qquad - \tfrac{1}{2}\,\mathbb{E}(|X - X'|\,|X - X''|) - \mathbb{E}(X|X - X'|)\\ &\quad= \tfrac{1}{4}\,\mathrm{dCov}^2(X, X) + \mathbb{E}(X^2) - \mathbb{E}(X|X - X'|).\end{aligned}$$

Finally,

$$\begin{aligned}\mathrm{dCov}^2(X, \varepsilon X) &= \mathbb{E}(|X - X'|\,|\varepsilon X - \varepsilon' X'|) + \mathbb{E}(|X - X'|)\mathbb{E}(|\varepsilon X - \varepsilon' X'|)\\ &\qquad - 2\mathbb{E}(|X - X'|\,|\varepsilon X - \varepsilon'' X''|)\\ &= \mathbb{E}(|X - X'|\,|X - \varepsilon' X'|) + \mathbb{E}(|X - X'|)\mathbb{E}(|X - \varepsilon' X'|)\\ &\qquad - 2\mathbb{E}(|X - X'|\,|X - \varepsilon'' X''|)\\ &= \tfrac{1}{2}\,\mathbb{E}\big((X - X')^2\big) + \tfrac{1}{2}\,\big(\mathbb{E}(|X - X'|)\big)^2 - \mathbb{E}(|X - X'|\,|X - X''|)\\ &= \tfrac{1}{2}\,\mathrm{dCov}^2(X, X).\end{aligned}$$

Consequently we have

$$\mathrm{dCor}(X, \varepsilon X) = \left(\frac{\mathrm{dCov}^2(X, X)}{\mathrm{dCov}^2(X, X) + 4\mathbb{E}(X^2) - 4\mathbb{E}(X|X - X'|)}\right)^{1/4}.$$

Let X be an indicator with $\mathbb{E}X = p$. Then $|X - X'|$ is an indicator as well, and

$$\begin{aligned}\mathbb{E}(|X - X'|) &= 2p(1 - p),\\ \mathbb{E}\big((X - X')^2\big) &= \mathbb{E}(|X - X'|) = 2p(1 - p),\\ \mathbb{E}(|X - X'|\,|X - X''|) &= p\big(\mathbb{E}(1 - X)\big)^2 + (1 - p)\big(\mathbb{E}X\big)^2\\ &= p(1 - p)^2 + (1 - p)p^2 = p(1 - p),\\ \mathbb{E}(X|X - X'|) &= p\mathbb{E}(1 - X) = p(1 - p),\end{aligned}$$

hence

$$\mathrm{dCor}(X, \varepsilon X) = \left(\frac{(1 - p)^2}{(1 - p)^2 + 1}\right)^{1/4}.$$

On the other hand, we will show that

$$\mathrm{dCov}^2(X, X) < 4\mathbb{E}(X^2) - 4\mathbb{E}(X|X - X'|)$$

for arbitrary nonnegative (and not identically zero) X, hence the supremum of the distance correlation coefficient between a symmetrically distributed random variable and its absolute value is $2^{-1/4}$.

Indeed, introduce $\varphi(t) = \mathbb{E}(|X - t|)$, then

$$\begin{aligned}\mathbb{E}(|X - X'|) &= \mathbb{E}\big(\varphi(X)\big),\\ \mathbb{E}(|X - X'|\,|X - X''|) &= \mathbb{E}\big(\varphi(X)^2\big),\\ \mathbb{E}(X|X - X'|) &= \mathbb{E}\big(X\varphi(X)\big),\end{aligned}$$

thus

$$
\begin{aligned}
&4\mathbb{E}(X^2) - 4\mathbb{E}(X|X-X'|) - \mathrm{dCov}^2(X,X)\\
&\quad = 4\mathbb{E}(X^2) - 4\mathbb{E}\big(X\varphi(X)\big) - 2\mathbb{E}(X^2) + 2\big(\mathbb{E}X\big)^2\\
&\qquad\quad - \big(\mathbb{E}(\varphi(X))\big)^2 + 2\mathbb{E}\big(\varphi(X)^2\big)\\
&\quad = 2\mathbb{E}\big((X-\varphi(X))^2\big) + 2\mathbb{E}(X^2) - \big(\mathbb{E}(\varphi(X))\big)^2\\
&\qquad\quad - 2\big(\mathbb{E}X - \mathbb{E}(\varphi(X))\big)^2 + \big(2\mathbb{E}X - \mathbb{E}(\varphi(X))\big)^2\\
&\quad = \mathrm{Var}\big(X-\varphi(X)\big) + \big(2\mathbb{E}X - \mathbb{E}(\varphi(X))\big)^2,
\end{aligned}
$$

and the right-hand side is strictly positive, because

$$\mathbb{E}(\varphi(X)) = \mathbb{E}(|X-X'|) < \mathbb{E}(X+X') = 2\mathbb{E}X. \quad \square$$

Is this the real supremum of the distance correlation coefficient for uncorrelated random variables? The answer is negative. In Section 3 we will show that the set of pairs $(\mathrm{cor}(X,Y),\ \mathrm{dCor}(X,Y))$ as X and Y run over the square integrable non-constant random variables contains the whole rectangle $(-1,1)\times(0,1)$.

3. Let us up the ante: α-distance correlation

Distance covariance can be generalized by taking the α-th powers of the distances in (1), with $0<\alpha<2$. That is, let

$$
\begin{aligned}
\mathrm{dCov}^2_\alpha(X,Y) = {}& \mathbb{E}\big(|X-X'|^\alpha|Y-Y'|^\alpha\big) + \mathbb{E}\big(|X-X'|^\alpha\big)\mathbb{E}\big(|Y-Y'|^\alpha\big)\\
&-2\mathbb{E}\big(|X-X'|^\alpha|Y-Y''|^\alpha\big).
\end{aligned}
\tag{4}
$$

By this definition, the α-distance correlation coefficient can be expressed in the usual way, as

$$\mathrm{dCor}_\alpha(X,Y) = \frac{\mathrm{dCov}_\alpha(X,Y)}{\sqrt{\mathrm{dCov}_\alpha(X,X)\mathrm{dCov}_\alpha(Y,Y)}}.$$

provided the denominator is positive. dCor_α shares all advantageous properties of dCor; in particular $\mathrm{dCor}_\alpha(X,Y)=0$ if and only if X and Y are independent (Székely and Rizzo, 2009). Moreover, while definition (4) only holds for random variables X and Y with finite moments of order 2α, the definition of $\mathrm{dCor}_\alpha(X,Y)$ can be straightforwardly extended to X,Y with moments of order α. For $\alpha=1$ we get back the distance correlation coefficient of (1).

Theorem 2. *Let $0<\alpha<2$. For every pair (ϱ,r), $-1<\varrho<1$, $0<r<1$, there exist random variables X,Y with finite moments of order 2α, such that $\mathrm{cor}(X,Y)=\varrho$, $\mathrm{dCor}_\alpha(X,Y)=r$.*

Remark. In addition to the set above, the only possible values of the pair $\big(\mathrm{cor}(X,Y),\ \mathrm{dCor}_\alpha(X,Y)\big)$ are $(-1,1)$, $(0,0)$ and $(1,1)$.

Proof. Let (U_1,U_2) have bivariate Rademacher distribution such that

$$
\begin{aligned}
&\mathbb{P}\big((U_1,U_2)=(1,1)\big) = \mathbb{P}\big((U_1,U_2)=(-1,-1)\big) = p/2,\\
&\mathbb{P}\big((U_1,U_2)=(-1,1)\big) = \mathbb{P}\big((U_1,U_2)=(1,-1)\big) = (1-p)/2.
\end{aligned}
$$

Similarly, let (V_1,V_2) have bivariate Rademacher distribution with

$$
\begin{aligned}
&\mathbb{P}\big((V_1,V_2)=(1,1)\big) = \mathbb{P}\big((V_1,V_2)=(-1,-1)\big) = q/2,\\
&\mathbb{P}\big((V_1,V_2)=(-1,1)\big) = \mathbb{P}\big((V_1,V_2)=(1,-1)\big) = (1-q)/2.
\end{aligned}
$$

Then U_1,U_2,V_1,V_2 are symmetric random variables. Let $0<\varepsilon<1$, $1/2<\delta<1/\alpha$, and define $(X_\varepsilon,Y_\varepsilon)$ as the mixture of (U_1,U_2) and $\varepsilon^\delta(V_1,V_2)$ with weights ε and $1-\varepsilon$, respectively. In other words, let Z_ε be a Bernoulli random variable with $\mathbb{P}(Z_\varepsilon=1)=1-\mathbb{P}(Z_\varepsilon=0)=\varepsilon$, and let Z_ε be independent of U_1,U_2,V_1,V_2, then we can write

$$(X_\varepsilon,Y_\varepsilon) = Z_\varepsilon(U_1,U_2) + (1-Z_\varepsilon)(V_1,V_2).$$

Elementary calculation yields

$$
\begin{aligned}
&\mathrm{cov}(X_\varepsilon,Y_\varepsilon) = \mathbb{E}(X_\varepsilon Y_\varepsilon) = \varepsilon(2p-1) + (1-\varepsilon)\varepsilon^{2\delta}(2q-1),\\
&\mathrm{Var}(X_\varepsilon) = \mathrm{Var}(Y_\varepsilon) = \mathbb{E}(X_\varepsilon^2) = \varepsilon + (1-\varepsilon)\varepsilon^{2\delta},
\end{aligned}
$$

hence

$$\mathrm{cor}(X_\varepsilon,Y_\varepsilon) = \frac{(2p-1)+(1-\varepsilon)\varepsilon^{2\delta-1}(2q-1)}{1+(1-\varepsilon)\varepsilon^{2\delta-1}}.$$

For fixed ε, δ and q try to choose p in such a way that $\mathrm{cor}(X_\varepsilon, Y_\varepsilon) = \varrho$; that is,

$$p = \tfrac{1}{2}\big[\varrho + 1 + (\varrho + 1 - 2q)(1-\varepsilon)\varepsilon^{2\delta-1}\big], \tag{5}$$

supposed $0 < p < 1$. If ε is sufficiently small, this will hold.

For $\mathrm{dCov}^2_\alpha(X_\varepsilon, Y_\varepsilon)$, let us start with the first term on the right-hand side of (4).

$$\begin{aligned}\mathbb{E}\big(|X_\varepsilon - X'_\varepsilon|^\alpha |Y_\varepsilon - Y'_\varepsilon|^\alpha\big) = {} & \varepsilon^2 \mathbb{E}\big(|U_1 - U'_1|^\alpha |U_2 - U'_2|^\alpha\big) \\ & + 2\varepsilon(1-\varepsilon)\mathbb{E}\big(|U_1 - \varepsilon^\delta V'_1|^\alpha |U_2 - \varepsilon^\delta V'_2|^\alpha\big) \\ & + (1-\varepsilon)^2 \varepsilon^{2\alpha\delta}\mathbb{E}\big(|V_1 - V'_1|^\alpha |V_2 - V'_2|^\alpha\big).\end{aligned}$$

All three expectations on the right-hand side are clearly bounded by $2^{2\alpha} < 16$. Furthermore,

$$\begin{aligned}\mathbb{E}\big(|U_1 - \varepsilon^\delta V'_1|^\alpha |U_2 - \varepsilon^\delta V'_2|^\alpha\big) = {} & \frac{pq + (1-p)(1-q)}{2}\big[(1-\varepsilon^\delta)^{2\alpha} + (1+\varepsilon^\delta)^{2\alpha}\big] \\ & + \frac{p(1-q) + (1-p)q}{2}(1-\varepsilon^{2\delta})^\alpha.\end{aligned}$$

Expanding the terms $(1-\varepsilon^\delta)^{2\alpha}$ and $(1+\varepsilon^\delta)^{2\alpha}$ into Maclaurin series yields

$$(1-\varepsilon^\delta)^{2\alpha} = 1 - 2\alpha\varepsilon^\delta + O(\varepsilon^{2\delta}), \quad (1+\varepsilon^\delta)^{2\alpha} = 1 + 2\alpha\varepsilon^\delta + O(\varepsilon^{2\delta}).$$

Hence $\Big|\mathbb{E}\big(|U_1 - \varepsilon^\delta V'_1|^\alpha |U_2 - \varepsilon^\delta V'_2|^\alpha\big) - 1\Big| = O(\varepsilon^{2\delta})$ and thus

$$\mathbb{E}\big(|X_\varepsilon - X'_\varepsilon|^\alpha |Y_\varepsilon - Y'_\varepsilon|^\alpha\big) = 2\varepsilon + \varepsilon^{2\alpha\delta}\mathbb{E}\big(|V_1 - V'_1|^\alpha |V_2 - V'_2|^\alpha\big) + R_1,$$

where the remainder R_1 satisfies $|R_1| \le \kappa_1\big(\varepsilon^{1+2\alpha\delta} + \varepsilon^2\big)$ with an absolute constant κ_1 not depending on ε, α and q.

Similarly, for the second term we can write

$$\begin{aligned}\mathbb{E}\big(|X_\varepsilon - X'_\varepsilon|^\alpha\big) &= \varepsilon^2\mathbb{E}\big(|U_1 - U'_1|^\alpha\big) + 2\varepsilon(1-\varepsilon)\mathbb{E}\big(|U_1 - \varepsilon^\delta V'_1|^\alpha\big) \\ &\quad + (1-\varepsilon)^2\varepsilon^{\alpha\delta}\mathbb{E}\big(|V_1 - V'_1|^\alpha\big) \\ &= \varepsilon^{\alpha\delta}\mathbb{E}\big(|V_1 - V'_1|^\alpha\big) + R,\end{aligned}$$

where $|R| \le 16\varepsilon$. Similar estimation holds for $\mathbb{E}\big(|Y_\varepsilon - Y'_\varepsilon|^\alpha\big)$. Hence

$$\mathbb{E}\big(|X_\varepsilon - X'_\varepsilon|^\alpha\big)\mathbb{E}\big(|Y_\varepsilon - Y'_\varepsilon|^\alpha\big) = \varepsilon^{2\alpha\delta}\mathbb{E}\big(|V_1 - V'_1|^\alpha\big)\mathbb{E}\big(|V_2 - V'_2|^\alpha\big) + R_2$$

with a remainder $|R_2| \le \kappa_2\,\varepsilon^{1+\alpha\delta}$, where the constant κ_2 is absolute again.

Finally, let us turn to the third term.

$$\begin{aligned}\mathbb{E}\big(|X_\varepsilon - X'_\varepsilon|^\alpha |Y_\varepsilon - Y''_\varepsilon|^\alpha\big) = {} & \varepsilon^3\mathbb{E}\big(|U_1 - U'_1|^\alpha |U_2 - U''_2|^\alpha\big) \\ & + \varepsilon^2(1-\varepsilon)\mathbb{E}\big(|\varepsilon^\delta V_1 - U'_1|^\alpha |\varepsilon^\delta V_2 - U''_2|^\alpha\big) \\ & + 2\varepsilon^2(1-\varepsilon)\mathbb{E}\big(|U_1 - \varepsilon^\delta V'_1|^\alpha |U_2 - U''_2|^\alpha\big) \\ & + \varepsilon(1-\varepsilon)^2\mathbb{E}\big(|U_1 - \varepsilon^\delta V'_1|^\alpha |U_2 - \varepsilon^\delta V''_2|^\alpha\big) \\ & + 2\varepsilon(1-\varepsilon)^2\varepsilon^{\alpha\delta}\mathbb{E}\big(|V_1 - V'_1|^\alpha |\varepsilon^\delta V_2 - U''_2|^\alpha\big) \\ & + (1-\varepsilon)^3\varepsilon^{2\alpha\delta}\mathbb{E}\big(|V_1 - V'_1|^\alpha |V_2 - V''_2|^\alpha\big).\end{aligned}$$

Similar computations as in the case of the first and the second terms yield

$$\mathbb{E}\big(|X_\varepsilon - X'_\varepsilon|^\alpha |Y_\varepsilon - Y''_\varepsilon|^\alpha\big) = \varepsilon + \varepsilon^{2\alpha\delta}\mathbb{E}\big(|V_1 - V'_1|\,|V_2 - V''_2|\big) + R_3,$$

where $|R_3| \le \kappa_3\,\varepsilon^{1+\alpha\delta}$ with an absolute constant κ_3. Indeed, the only nontrivial term is

$$\mathbb{E}\big(|U_1 - \varepsilon^\delta V'_1|^\alpha |U_2 - \varepsilon^\delta V''_2|^\alpha\big) = \frac{1}{2}\left[\frac{(1-\varepsilon^\delta)^{2\alpha} + (1+\varepsilon^\delta)^{2\alpha}}{2} + (1-\varepsilon^{2\delta})^\alpha\right],$$

which we have already shown to be $1 + O\big(\varepsilon^{2\delta}\big)$.

Combining the above estimations, we find the terms of order ε cancel out yielding

$$\begin{aligned}\mathrm{dCov}^2_\alpha(X_\varepsilon, Y_\varepsilon) &= \varepsilon^{2\alpha\delta}\mathrm{dCov}^2_\alpha(V_1, V_2) + O\big(\varepsilon^{1+\alpha\delta}\big) \\ &= \varepsilon^{2\alpha\delta}2^{2\alpha-2}(2q-1)^2 + O\big(\varepsilon^{1+\alpha\delta}\big),\end{aligned}$$

where the constant involved in the big oh notation is absolute.

In the calculations above nothing was used about the actual values of p and q. Taking $p = q = 1$ we get that

$$\begin{aligned}\mathrm{dCov}_\alpha^2(X_\varepsilon, X_\varepsilon) = \mathrm{dCov}_\alpha^2(Y_\varepsilon, Y_\varepsilon) &= \varepsilon^{2\alpha\delta}\mathrm{dCov}_\alpha^2(V_1, V_2) + O\big(\varepsilon^{1+\alpha\delta}\big)\\ &= \varepsilon^{2\alpha\delta}2^{2\alpha-2} + O\big(\varepsilon^{1+\alpha\delta}\big).\end{aligned}$$

Therefore we obtain

$$\mathrm{dCor}_\alpha(X_\varepsilon, Y_\varepsilon) = |2q-1| + O\big(\varepsilon^{1-\alpha\delta}\big),$$

with an absolute constant in the big oh. This will be important if we want ε or q to vary, because then p also varies according to (5).

Let $0 < r_1 < r < r_2$ and $q_1 = (r_1+1)/2$, $q_2 = (r_2+1)/2$. Let ε be sufficiently small so that p_1 and p_2 given by (5) fall between 0 and 1, furthermore $\mathrm{dCor}_\alpha(X_\varepsilon, Y_\varepsilon) < r$ for $q = q_1$ and $\mathrm{dCor}_\alpha(X_\varepsilon, Y_\varepsilon) > r$ for $q = q_2$. Let us move q from q_1 to q_2. It is easy to see that $\mathrm{dCor}_\alpha(X_\varepsilon, Y_\varepsilon)$ continuously depends on q, hence there must exist a q between q_1 and q_2 for which $\mathrm{dCor}_\alpha(X_\varepsilon, Y_\varepsilon) = r$ (and, of course, $\mathrm{cor}(X_\varepsilon, Y_\varepsilon) = \varrho$.) $\square$

Conjecture. *Let $0 < \alpha < \beta \le 2$. Then for every pair $0 < \varrho, r < 1$ there exist random variables X, Y with finite moments of order α and β, resp., such that $\mathrm{dCor}_\alpha(X, Y) = \varrho$ and $\mathrm{dCor}_\beta(X, Y) = r$.*

CRediT authorship contribution statement

Dominic Edelmann: Conceptualization, Formal analysis, Writing - original draft, Writing - Review & Editing. **Tamás F. Móri:** Conceptualization, Formal analysis, Writing - original draft, Writing - Review & Editing. **Gábor J. Székely:** Conceptualization, Formal analysis, Writing - original draft, Writing - Review & Editing.

Acknowledgments

Funding

This research was supported by the Deutsche Forschungsgemeinschaft (DFG), Germany [project number 417754611, to D. Edelmann]; and by the Hungarian National Research, Development and Innovation Office NKFIH [Grant No. K125569, to T. F. Móri].

Part of this research was supported by the National Science Foundation, USA while the third author was working at the Foundation.

References

Dueck, J., Edelmann, D., Richards, D., 2017. Distance correlation coefficients for Lancaster distributions. J. Multivar. Anal. 154, 19–39. http://dx.doi.org/10.1016/j.jmva.2016.10.012.

Székely, G.J., Rizzo, M.L., 2009. Brownian distance covariance. Ann. Appl. Stat. 3, 1236–1265. http://dx.doi.org/10.1214/09-AOAS312.

Székely, G.J., Rizzo, M.L., Bakirov, N.K., 2007. Measuring and testing independence by correlation of distances. Ann. Stat. 35, 2769–2794. http://dx.doi.org/10.1214/009053607000000505.