Electronic Journal of Statistics
Vol. 18 (2024) 653–673
ISSN: 1935-7524
https://doi.org/10.1214/24-EJS2212

A note on the equivalence between the conditional uncorrelation and the independence of random variables

Piotr Jaworski[1], **Damian Jelito**[2] **and Marcin Pitera**[2]

[1] *Institute of Mathematics, University of Warsaw, Warsaw, Poland,e-mail:* p.jaworski@mimuw.edu.pl

[2] *Institute of Mathematics, Jagiellonian University, Kraków, Poland,e-mail:* damian.jelito@uj.edu.pl*;* marcin.pitera@uj.edu.pl

Abstract: It is well known that while the independence of random variables implies zero correlation, the opposite is not true. Namely, uncorrelated random variables are not necessarily independent. In this note we show that the implication could be reversed if we consider the localised version of the correlation coefficient. More specifically, we show that if random variables are conditionally (locally) uncorrelated for any quantile conditioning sets, then they are independent. For simplicity, we focus on the absolutely continuous case. Also, we illustrate potential usefulness of the stated result using multiple examples.

MSC2020 subject classifications: Primary 60E05, 62E10, 62H20.
Keywords and phrases: Correlation, Pearson's correlation, Linear dependence, Zero conditional correlation, Zero conditional covariance, Independence, Linear independence, Local correlation.

Received May 2023.

1. Introduction

The concept of linear correlation was first presented in [12], see [30] for a historical note on the correlation invention. While mathematically simple and elegant, statistical analysis based on correlation measurement could be confusing and lead to subtle errors if not treated with caution, see e.g. [3], [34], and references therein. Since correlation aims to measure the linear dependence between random variables, it often fails to properly capture non-linear structures. Although the dependence could be fully described using the copula function, it is more appealing, especially to practitioners, to use simpler (numeric) characteristics to describe the degree of dependence, see [23] or [22]. Because of that, a lot of alternative measures of dependence have been proposed in the literature and this field is constantly evolving. Let us alone mention the concepts of concordance measures, entropy correlations, projection correlations, tail correlations, partial and conditional correlations, maximal correlations, time-varying dynamic

arXiv: 2210.16655

correlations, local Gaussian correlations, and distance correlations based on energy statistics, see [28, 27, 36, 2, 4, 13, 21, 35, 1, 32, 31, 33], and references therein.

Typically, it is expected that the zero value of a given dependence measure should, in some sense, imply independence. What is interesting, at first, the concept of null linear correlation was often mixed with independence and it took some time for statisticians to distinguish between null correlation and statistical independence, see [9]. Of course, it is currently well known that while the independence of random variables implies zero correlation, the opposite is not true, see e.g. [6] for a classroom example.

In this short paper we answer a simple question about how one can revert the aforementioned implication, i.e. whether one can use linear correlation to study (proper) independence. Allowing non-linear transforms of random variables, the reverse implication is in fact trivially true as one of the alternative definition of independence states that two random variables X and Y are independent, if $f(X)$ and $g(X)$ are uncorrelated for any test functions f and g; in fact, it is sufficient to consider set indicator functions to directly recover the definition of independence. Still, this characterisation is not appealing from practical perspective since it is hard to pre-set the family of test functions that would work for any arbitrary pair of random variables and lead to efficient statistical setup. Another approach is to consider a localised version of correlation and study its properties, see Section 6 in [22] for details. In this paper, following [18], we propose to bind those two approaches together and consider a family of conditional correlations, where the conditioning is based on the quantile set linked to the values of X and Y, see Section 3 for details. In the main result of this paper, Theorem 3.1, we show that null correlation on every quantile set implies independence of random variables so that the aforementioned implication could be reverted by looking locally into linear relation between random variables. Namely, we show that random variables are independent if and only if they are locally linearly independent. It is worth noting that a similar results is true in the multivariate case, i.e. conditional correlation matrices could be used to characterise mutual independence, see Theorem 3.2. Due to our best knowledge, those results have not been directly analysed previously in the literature – this is most likely due to the fact that localised correlations considered so far were not bound directly to quantile sets allowing efficient local treatment. Also, note that this paper effectively provides a solution to the Advanced Problem 6327 in [7]; the problem is listed as open therein, see also [26].

We believe that our proposal could be appealing to practitioners and could lead to development of new efficient statistical frameworks. In fact, the sample version of (local) quantile correlation could be easily computed using rank statistics and exhibits statistical properties similar to the unconditional correlation; this aspect is left to future research. In other words, the results presented in this paper lay the theoretical ground to expansion of the statistical framework based on quantile conditional moments which already proved to be useful, see e.g. [16], [19], and [25]. As an example, one could define the conditional version of the auto-correlation function that could be used to study time-series which exhibits heavy tails, see Example 3 for details, or study the tail-based correlations to recover dependence conditioned on tail-events, see [17].

This paper is organised as follows. In Section 2, we introduce the basic notation and define the concept of quantile conditional correlation. In Section 3, we state and prove the main result, Theorem 3.1, together with its multidimensional extensions. Next, in Section 4, we discuss potential applications and provide examples that illustrate how our approach could be used to study dependence between random variables. Finally, in Section 5 we provide concluding remarks.

2. Preliminaries

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let (X, Y) be a random vector defined on this space. By Sklar's theorem, we know that the joint distribution of (X, Y) can be represented as

$$\mathbb{P}[X \leq x, Y \leq y] = C(F_X(x), F_Y(y)), \quad x, y \in \mathbb{R}, \tag{1}$$

where F_X and F_Y denote the distributions of X and Y, respectively, and C is the copula function of the vector (X, Y), see e.g. Theorem 2.3.3 in [23]. For simplicity, from now on we assume that the vector (X, Y) is absolutely continuous and use f_X, f_Y, f, and c, to denote the density functions of X, Y, (X, Y), and C, respectively. Also, we assume that F_X and F_Y are bijective, the vector (X, Y) has a full (non-degenerate) support, and the copula density c is continuous. In this case, the copula function C is unique and can be easily recovered from the joint distribution using the formula $C(u, v) = \mathbb{P}[X \leq Q_X(u), Y \leq Q_Y(v)]$, $u, v \in (0, 1)$, where $Q_X := F_X^{-1}$ and $Q_Y := F_Y^{-1}$ are the quantile functions of X and Y, respectively.

Now, let us introduce a notation associated with quantile conditional covariances. Given a set $A \in \mathcal{F}$, we define the conditional covariance of (X, Y) on A by setting

$$\mathrm{Cov}_A[X, Y] := \mathbb{E}\left[XY|A\right] - \mathbb{E}\left[X|A\right]\mathbb{E}\left[Y|A\right], \tag{2}$$

provided that the expectations are well-defined. In this paper, we are interested in quantile-based conditioning. Namely, given a vector (X, Y) and *quantile splits* $0 < p_1 < q_1 < 1$ and $0 < p_2 < q_2 < 1$, we define the corresponding *quantile set* as

$$A := \{\omega \in \Omega \colon Q_X(p_1) \leq X(\omega) \leq Q_X(q_1), Q_Y(p_2) \leq Y(\omega) \leq Q_Y(q_2)\}. \tag{3}$$

Note that since we assumed a full support, we get $\mathbb{P}[A] > 0$ for any quantile split. Also, since both X and Y are bounded on A, we get that (2) is well-defined and finite. Thus, we can also define the corresponding conditional correlation by setting

$$\mathrm{Cor}_A[X, Y] := \frac{\mathrm{Cov}_A[X, Y]}{\sqrt{\mathrm{Var}_A[X]\,\mathrm{Var}_A[Y]}}, \tag{4}$$

where $\mathrm{Var}_A[X] := \mathbb{E}[X^2|A] - \mathbb{E}^2[X|A]$ and $\mathrm{Var}_A[Y] := \mathbb{E}[Y^2|A] - \mathbb{E}^2[Y|A]$ are conditional covariances of X and Y, respectively.

From now on we assume that we are given specific quantile splits $0 < p_1 < q_1 < 1$ and $0 < p_2 < q_2 < 1$, and use A to denote the corresponding quantile set as defined in (3). For brevity, we also introduce the corresponding value projection set $\tilde{A} := [Q_X(p_1), Q_X(q_1)] \times [Q_Y(p_2), Q_Y(q_2)] \subset \mathbb{R}^2$. With this notation, we get that (2) could be expressed as

$$\mathrm{Cov}_A[X,Y] = \frac{1}{\mathbb{P}[A]} \int_{\tilde{A}} xyf(x,y)dydx - \frac{1}{\mathbb{P}^2[A]} \int_{\tilde{A}} xf(x,y)dydx \int_{\tilde{A}} yf(x,y)dydx. \quad (5)$$

We say that X and Y are conditionally uncorrelated on A, if $\mathrm{Cor}_A[X,Y] = 0$.

3. Main result

In this section we present the main result of this note, which shows that the independence of random variables could be linked to their conditional uncorrelation on any quantile set.

Theorem 3.1. *Random variables X and Y are independent if and only if they are conditionally uncorrelated on every quantile set, i.e. for any quantile splits $0 < p_1 < q_1 < 1$ and $0 < p_2 < q_2 < 1$ and the related set A, we get*

$$\mathrm{Cor}_A[X,Y] = 0.$$

Proof. The fact that the independence of X and Y implies $\mathrm{Cor}_A[X,Y] = 0$ for any $A \in \mathcal{F}$ follows from the standard argument which is omitted for brevity. Let us now assume that for any quantile splits $0 < p_1 < q_1 < 1$ and $0 < p_2 < q_2 < 1$ and the related set A we get $\mathrm{Cor}_A[X,Y] = 0$ or equivalently

$$\mathrm{Cov}_A[X,Y] = 0. \quad (6)$$

First, let us show that for any $u_1, v_1, u_2, v_2 \in (0,1)$ we have

$$c(u_1,v_1)c(u_2,v_2) - c(u_1,v_2)c(u_2,v_1) = 0. \quad (7)$$

We start with deriving a useful representation of $\mathrm{Cov}_A[X,Y]$ based on (5). First, define $V_c(u_1,v_1,u_2,v_2) := c(u_1,v_1)c(u_2,v_2) - c(u_1,v_2)c(u_2,v_1)$, $u_1,v_1,u_2,v_2 \in (0,1)$, and note that V_c is anti-symmetric in (u_1,u_2) and (v_1,v_2), i.e. we get $V_c(u_2,v_1,u_1,v_2) = -V_c(u_1,v_1,u_2,v_2)$ and $V_c(u_1,v_2,u_2,v_1) = -V_c(u_1,v_1,u_2,v_2)$. Next, for any set A defined in (3), using the fact that

$$f(x,y) = c(F_X(x),F_Y(y))f_X(x)f_Y(y), \quad x,y \in \mathbb{R},$$

and substituting $x = Q_X(u)$ and $y = Q_Y(v)$, we get

$$\mathrm{Cov}_A[X,Y] = \frac{1}{\mathbb{P}[A]} \int_{p_1}^{q_1} \int_{p_2}^{q_2} Q_X(u)Q_Y(v)c(u,v)dvdu$$

$$
\begin{aligned}
&- \frac{1}{\mathbb{P}^2[A]} \int_{p_1}^{q_1} \int_{p_2}^{q_2} Q_X(u) c(u,v) dv du \int_{p_1}^{q_1} \int_{p_2}^{q_2} Q_Y(v) c(u,v) dv du \\
&= \frac{1}{\mathbb{P}^2[A]} \int_{p_1}^{q_1} \int_{p_2}^{q_2} \int_{p_1}^{q_1} \int_{p_2}^{q_2} Q_X(u_1) Q_Y(v_1) \times \\
&\qquad \times V_c(u_1, v_1, u_2, v_2) dv_2 du_2 dv_1 du_1. \qquad (8)
\end{aligned}
$$

For any $u_1, v_1, u_2, v_2 \in (0,1)$, let us define

$$
H(u_1, v_1, u_2, v_2) := (Q_X(u_1) - Q_X(u_2))(Q_Y(v_1) - Q_Y(v_2)) V_c(u_1, v_1, u_2, v_2).
$$

Thus, using (6) and the anti-symmetry of V_c, we get that (8) implies

$$
0 = \int_{p_1}^{q_1} \int_{p_2}^{q_2} \int_{p_1}^{q_1} \int_{p_2}^{q_2} H(u_1, v_1, u_2, v_2) dv_2 du_2 dv_1 du_1. \qquad (9)
$$

Next, using the multi-variable chain rule to differentiate with respect to q_1 and changing the order of integration, we get

$$
\begin{aligned}
0 = &\int_{p_1}^{q_1} \int_{p_2}^{q_2} \int_{p_2}^{q_2} H(q_1, v_1, u_2, v_2) dv_2 dv_1 du_2 \\
&\qquad + \int_{p_1}^{q_1} \int_{p_2}^{q_2} \int_{p_2}^{q_2} H(u_1, v_1, q_1, v_2) dv_2 dv_1 du_1,
\end{aligned}
$$

and consequently, due to the symmetry of H (in (u_1, u_2)), we have

$$
0 = \int_{p_1}^{q_1} \int_{p_2}^{q_2} \int_{p_2}^{q_2} H(q_1, v_1, u_2, v_2) dv_2 dv_1 du_2. \qquad (10)
$$

Thus, differentiating (10) with respect to p_1 yields

$$
0 = \int_{p_2}^{q_2} \int_{p_2}^{q_2} H(q_1, v_1, p_1, v_2) dv_2 dv_1.
$$

Performing a similar operation again, i.e. differentiating with respect to q_2 and then p_2, we finally get

$$
H(q_1, q_2, p_1, p_2) = 0. \qquad (11)
$$

Now, using the strict monotonicity of Q_X and Q_Y, we get

$$
(Q_X(q_1) - Q_X(p_1))(Q_Y(q_2) - Q_Y(p_2)) > 0.
$$

Thus, directly from the definition of H, we get that (11) implies

$$
c(q_1, q_2) c(p_1, p_2) - c(q_1, p_2) c(p_1, q_2) = 0, \qquad (12)
$$

for $0 < p_1 < q_1 < 1$ and $0 < p_2 < q_2 < 1$. This concludes the proof of (7) since for $q_1 = p_1$ or $q_2 = p_2$ the equality (12) is trivial, and the symmetry of H allows us to easily extend (11) to the full parameter space $q_1, q_2, p_1, p_2 \in (0,1)$.

Second, let us show that (7) implies the independence of X and Y. Using (7) and recalling that C could be seen as a distribution function of a random vector with marginals distributed uniformly on $[0,1]$, for any $x, y \in [0,1]$ we get

$$\begin{aligned} C(x,y) &= \int_0^x \int_0^y c(u,v)dvdu = \int_0^x \int_0^y \int_0^1 \int_0^1 c(u_1,v_1)c(u_2,v_2)dv_2du_2dv_1du_1 \\ &= \int_0^x \int_0^y \int_0^1 \int_0^1 c(u_1,v_2)c(u_2,v_1)dv_2du_2dv_1du_1 \\ &= \int_0^x \int_0^1 c(u_1,v_2)dv_2du_1 \int_0^1 \int_0^y c(u_2,v_1)dv_1du_2 \\ &= C(x,1)C(1,y) \\ &= xy, \end{aligned} \tag{13}$$

which shows that the copula of (X,Y) is the product copula. Recalling (1), we get that X and Y are independent, which concludes the proof. □

Remark 1 (Conditional Spearman's ρ and independence)**.** From Theorem 3.1 one can easily deduce that random variables X and Y are independent if and only if conditional Spearman's ρ coefficient on every quantile set is equal to zero. To prove this it is enough to observe that Spearman's ρ is in fact Pearson's correlation applied to the copula function.

Remark 2 (Local linear independence implies independence)**.** By investigating the proof of Theorem 3.1 one can see that proving Equality (12) is a key step in establishing independence. While in Theorem 3.1 we did not set any restriction on quantile split values $0 < p_1 < q_1 < 1$ and $0 < p_1 < q_1 < 1$, it is in fact sufficient to require that for any quantile point $(Q_X(p), Q_Y(q))$, $p, q \in (0,1)$, the quantile conditional correlations are null inside some neighbourhood of $(Q_X(p), Q_Y(q))$, e.g. for some $\epsilon > 0$ and any $p_1, q_1 \in (Q_X(p-\epsilon), Q_X(p+\epsilon))$ and $p_2, q_2 \in (Q_Y(q-\epsilon), Q_Y(q+\epsilon))$. Indeed, this implies that (9) is satisfied for any sufficiently small hypercubes $[p_1,q_1] \times [p_1,q_1] \times [p_2,q_2] \times [p_2,q_2]$ which can be combined to recover (9) for any $0 < p_1 < q_1 < 1$ and $0 < p_1 < q_1 < 1$ and, consequently, get Equality (12). This effectively shows than random variables X and Y are independent if and only if they are locally linearly independent.

Remark 3 (Tail-event dependence and spatial contagion)**.** From Theorem 3.1 we can see that to reject the (global) independence of X and Y, it is enough to find a single quantile split $0 < p_1 < q_1 < 1$ and $0 < p_2 < q_2 < 1$ on which the conditional quantile correlation is not equal to zero. In signal processing or financial time-series modelling, it is natural to consider left tail events, e.g. when one or both of the values q_1 and q_2 are small. Such events could be linked to the presence of the so-called *spatial contagion* in which dependence increases in the presence of system turbulence. This might be used to construct statistical frameworks based on quantile tail-event analysis, see [10], [17], and references therein.

Remark 4 (Conditional independence)**.** Theorem 3.1 statement could be also transferred to the conditional independence setting. Intuitively speaking, random vector (X, Y) has conditionally independent margins given a third random variable Z if and only if its Z-conditioned laws have independent margins, i.e. if condition $\mathbb{P}[X \leq x, Y \leq y | Z = z] = \mathbb{P}[X \leq x | Z = z]\mathbb{P}[Y \leq y | Z = z]$ holds for any $x, y, z \in \mathbb{R}$. We can also measure the conditional independence of (X,Y) given a specific value of Z and consider the respective conditional correlations – Theorem 3.1 ensures that all conditioned quantile correlations are null if and only if conditional independence holds.

As we show now, Theorem 3.1 could be extended to the multivariate case. To get this extension, we use two alternative approaches. First, in Theorem 3.2, we consider a conditional correlation matrix. Second, in Theorem 3.3 and Theorem 3.4, we use linear combinations of margins.

Before we state the result, let us introduce some notation. Consider an n-dimensional random vector $X = (X_1, \ldots, X_n)$ and assume that it satisfies the assumptions analogous to the ones used in Theorem 3.1, i.e. bijectiveness of the the marginal distribution functions, full support condition, absolute continuity of the joint distribution, and continuity of the copula density. Also, for any quantile splits $0 < p_i < q_i < 1$, $i = 1, \ldots, n$, we define the quantile set corresponding to $(X_1, \ldots, X_n)$ by

$$A := \bigcap_{i=1}^{n} \{Q_i(p_i) \leq X_i \leq Q_i(q_i)\}, \tag{14}$$

where Q_i is the quantile function of X_i, $i = 1, \ldots, n$. Finally, by Σ_A we denote the associated conditional correlation matrix with the entries given by $\Sigma_A[i,j] := [\mathrm{Cor}_A(X_i, X_j)]$, $i, j = 1, \ldots, n$; we also use I_n to denote the $n \times n$ identity matrix.

Theorem 3.2. *The n-dimensional random vector X has mutually independent margins if and only if, for any quantile splits $0 < p_i < q_i < 1$, $i = 1, \ldots, n$, and the related set A, its conditional correlation matrix Σ_A is equal to the identity matrix.*

Proof. The argument is based on the proof of Theorem 3.1 and we provide only an outline. Also, for simplicity, we consider only $n = 3$; the general case follows the same logic. As before, it is straightforward to check that the independence of margins imply diagonal conditional correlation matrix, so we focus on the reverse implication.

For simplicity and with a slight abuse of notation, we use C and c to denote the copula and the copula density corresponding to (X_1, X_2, X_3), respectively. Let us assume that for any quantile splits $0 < p_i < q_i < 1$, $i = 1, 2, 3$, and the related set A we get $\Sigma_A = \mathrm{I}_3$ or equivalently

$$\mathrm{Cov}_A[X_i, Y_j] = 0, \quad i, j \in \{1, 2, 3\},\ i \neq j. \tag{15}$$

Next, as in (8), we get

$$\mathrm{Cov}_A[X_1, X_2] = \frac{1}{2\mathbb{P}^2[A]} \int_{p_1}^{q_1} \int_{p_2}^{q_2} \int_{p_3}^{q_3} \int_{p_1}^{q_1} \int_{p_2}^{q_2} \int_{p_3}^{q_3} Q_X(u_1) Q_Y(v_1) \times \\ \times V_c(u_1, v_1, w_1, u_2, v_2, w_2) dw_2 dv_2 du_2 dw_1 dv_1 du_1, \quad (16)$$

where, for any $u_1, v_1, w_1, u_2, v_2, w_2 \in (0,1)$, we define

$$V_c(u_1, v_1, w_1, u_2, v_2, w_2) := c(u_1, v_1, w_1)c(u_2, v_2, w_2) - c(u_1, v_2, w_1)c(u_2, v_1, w_2).$$

Also, setting $Q(u_1, v_1, u_2, v_2) := (Q_1(u_1) - Q_1(u_2))(Q_2(v_1) - Q_2(v_2))$ for any $u_1, u_2, v_1, v_2 \in (0,1)$, and repeating the argument leading to (11), we get

$$0 = \int_{p_3}^{q_3} \int_{p_3}^{q_3} Q(q_1, q_2, p_1, p_2) V_c(q_1, q_2, w_1, p_1, p_2, w_2) dw_2 dw_1.$$

Noting that $Q(q_1, q_2, p_1, p_2) > 0$ and differentiating the iterated integral with respect to q_3 and p_3, for any $0 < p_i < q_i < 1$, $i = 1, 2, 3$, we get

$$\begin{aligned} 0 &= V_c(q_1, q_2, q_3, p_1, p_2, p_3) + V_c(q_1, q_2, p_3, p_1, p_2, q_3) \\ &= c(q_1, q_2, q_3)c(p_1, p_2, p_3) - c(q_1, p_2, q_3)c(p_1, q_2, p_3) \\ &\quad + c(q_1, q_2, p_3)c(p_1, p_2, q_3) - c(q_1, p_2, p_3)c(p_1, q_2, q_3). \end{aligned}$$

In fact, as in the proof of Theorem 3.1, we get that the formula is valid for any $p_i, q_i \in (0,1)$, $i = 1, 2, 3$; see the discussion following (12) for details. Using this observation, for any $x, y, z \in [0,1]$, as in (13), we get

$$\begin{aligned} &C(x, y, z) \\ &= \int_0^x \int_0^y \int_0^z \int_0^1 \int_0^1 \int_0^1 c(u_1, v_1, w_1)c(u_2, v_2, w_2) dw_2 dv_2 du_2 dw_1 dv_1 du_1 \\ &= \int_0^x \int_0^y \int_0^z \int_0^1 \int_0^1 \int_0^1 c(u_1, v_2, w_1)c(u_2, v_1, w_2) dw_2 dv_2 du_2 dw_1 dv_1 du_1 \\ &\quad - \int_0^x \int_0^y \int_0^z \int_0^1 \int_0^1 \int_0^1 c(u_1, v_1, w_2)c(u_2, v_2, w_1) dw_2 dv_2 du_2 dw_1 dv_1 du_1 \\ &\quad + \int_0^x \int_0^y \int_0^z \int_0^1 \int_0^1 \int_0^1 c(u_2, v_1, w_2)c(u_1, v_2, w_1) dw_2 dv_2 du_2 dw_1 dv_1 du_1 \\ &= C(x, 1, z)C(1, y, 1) - C(x, y, 1)C(1, 1, z) + C(1, y, z)C(x, 1, 1) \\ &= C(x, 1, z)y - C(x, y, 1)z + C(1, y, z)x. \end{aligned} \quad (17)$$

In particular, setting $z = 1$, we get $C(x, y, 1) = xy$, $x, z \in [0,1]$. Using the same argument applied to $\mathrm{Cov}_A[X_1, X_3]$ and $\mathrm{Cov}_A[X_2, X_3]$, we also get $C(x, 1, z) = xz$ and $C(1, y, z) = yz$, $x, y, z \in [0,1]$. Consequently, from (17), we get

$$C(x, y, z) = xyz, \quad x, y, z \in [0,1],$$

which concludes the proof. □

The next generalisation of Theorem 3.1 is based on linear combinations. For simplicity, given an n-dimentional random vector X and an m-dimensional random vector Y, we pre-assume that for any $\alpha \in \mathbb{R}^n \setminus \{0\}$ and $\beta \in \mathbb{R}^m \setminus \{0\}$, the random vector $(\langle X, \alpha\rangle, \langle Y, \beta\rangle)$, where $\langle \cdot, \cdot\rangle$ denotes the standard Euclidean inner product, satisfy our usual assumptions, i.e. bijectiveness of the the marginal distribution functions, full support condition, absolute continuity of the joint distribution, and continuity of the copula density.

Theorem 3.3. *Let X and Y be n-dimensional and m-dimensional random vectors, respectively. Then, X and Y are independent if and only if for any $\alpha \in \mathbb{R}^n \setminus \{0\}$ and $\beta \in \mathbb{R}^m \setminus \{0\}$, the random variables $\langle X, \alpha\rangle$ and $\langle Y, \beta\rangle$ are conditionally uncorrelated, i.e. for any quantile splits $0 < p_1 < q_1 < 1$ and $0 < p_2 < q_2 < 1$ and the related set A, defined for $(\langle X, \alpha\rangle, \langle Y, \beta\rangle)$, we get*

$$\mathrm{Cor}_A\left[\langle X, \alpha\rangle, \langle Y, \beta\rangle\right] = 0. \tag{18}$$

Proof. As in the proof of Theorem 3.1, we focus on the argument that (18) implies independence; the reverse implication is standard. Note that (18) combined with Theorem 3.1 implies that the random variables $\langle X, \alpha\rangle$ and $\langle Y, \beta\rangle$ are independent for any $\alpha \in \mathbb{R}^n$ and $\beta \in \mathbb{R}^m$. In particular, we get

$$\phi_{(X,Y)}(\alpha, \beta) = \phi_X(\alpha)\phi_Y(\beta), \quad \alpha \in \mathbb{R}^n,\ \beta \in \mathbb{R}^m, \tag{19}$$

where $\phi_Z(t) := \mathbb{E}[\exp(i\langle Z, t\rangle)]$, $t \in \mathbb{R}^d$, denotes the characteristic function of an arbitrary d-dimensional random vector Z. Combining (19) with Theorem 4, Section II.12, in [29] we conclude the proof. □

Theorem 3.3 can be used to get another characterisation of random vector margins independence based on a recursive scheme. For an n-dimensional vector $a := (a_1, \ldots, a_n)$ and $k = 1, \ldots, n$, let $a^{1:k}$ denote its subvector $a^{1:k} := (a_1, \ldots, a_k)$. Again, for simplicity, for n-dimensional random vector X, we assume that for any $k = 1, \ldots, n-1$ and $\alpha^k \in \mathbb{R}^k \setminus \{0\}$, the random variables X_{k+1} and $\langle X^{1:k}, \alpha^k\rangle$ satisfy our standard assumptions.

Theorem 3.4. *Let X be an n-dimensional random vector. Assume that for any $k = 1, \ldots, n-1$ and $\alpha^k \in \mathbb{R}^k \setminus \{0\}$, the random variables X_{k+1} and $\langle X^{1:k}, \alpha^k\rangle$ are conditionally uncorrelated. Then, the margins of X are mutually independent.*

Proof. Using Theorem 3.3, we get that, for any $\alpha^{n-1} \in \mathbb{R}^{n-1} \setminus \{0\}$, the random variables X_n and $\langle X^{1:(n-1)}, \alpha^{n-1}\rangle$ are independent. Hence, the characteristic function of the random vector X satisfies $\phi_X(\alpha) = \phi_{X_n}(\alpha_n)\phi_{X^{1:(n-1)}}(\alpha^{1:(n-1)})$, $\alpha \in \mathbb{R}^n$. In fact, inductively, we get that the characteristic funcion of X factorises into the product of the characteristic functions of the margins, i.e.

$$\phi_X(\alpha) = \prod_{k=1}^{n} \phi_{X_i}(\alpha_i), \quad \alpha \in \mathbb{R}^n.$$

Using Theorem 4, Section II.12, in [29] we conclude the proof. □

Remark 5 (Mutual independence for infinite series of random variables)**.** It should be noted that Theorem 3.4 could be used to get a characterisation of infinite series of random variables, e.g. $(X_t)_{t\in\mathbb{N}}$. Indeed, it is enough to recall that the mutual independence of the family $(X_t)_{t\in\mathbb{N}}$ means mutual independence of any finite subfamily of random variables.

4. Discussion on potential applications and examples

In this section we present five examples which show how conditional correlation analysis could be applied to study independence in various mathematical and statistical contexts. In particular, while the development of rigorous statistical frameworks is out of scope of this paper, we want to explain how the independence characterisation presented in Theorem 3.1 could be used to refine existing methods or to develop new ones. Before we present the examples, let us provide some generic remarks linked to potential statistical applications of conditional correlations.

First, it should be emphasized that statistical independence is an important aspect of many models and there are numerous tools designed to study it. In the following examples we focus on five distinct applications linked to: (1) identification of non-linear dependence structures; (2) generic statistical independence tests; (3) serial dependence tests; (4) verification of independence assumptions in regression models; (5) conditional and partial independence characterisation. That saying, we want to emphasize that this does not exhaust all possible applications – due to generic nature of Theorem 3.1 and Theorem 3.2, the analysis based on conditional correlations could be applied in almost any framework in which statistical independence is being evaluated. To name a few further examples, this could correspond to copula goodness-of-fit tests based on Rosenblatt transform, GARCH time-series assumption checks or the extension of the measure of association normative approach to local (conditional) coefficients.

Second, in all examples we estimate local correlation using a simple empirical estimator following the logic from [19]. Namely, given quantile splits $0 < p_1 < q_1 < 1$ and $0 < p_2 < q_2 < 1$, the corresponding set A, and a sample $\{(x_i, y_i)\}_{i=1}^{n}$ of size $n \in \mathbb{N}$, we compute conditional correlation using a simple two-step procedure. In the first step we determine the conditioned subsample by picking only observations which satisfy (empirical) rank constraints on each margin, see Figure 5 for an illustration. In the second step we compute the empirical correlation for the selected (conditioned) observation. It is relatively easy to show that the resulting empirical conditional correlation estimators will be consistent and will have CLT-type law, i.e. their asymptotic distribution will be normal. We refer to [19] for more details.

Third, we note that the statistical test of (full) independence is by design a very hard task. Indeed, one can imagine a situation of random variables that are dependent only on some set that has a very small probability. In such instance, one would need (statistically meaningful) information from this set which require methods that are sensitive to any local deviations from independence.

While this task could be achieved using conditional correlations and adaptive quantile split grids with grid decreasing diameter depending on the sample size, cf. Remark 2, we decided to focus on tests which detect dependence by extracting specific features from the sample. In this context, the flexibility coming from the freedom of quantile split choice allows us to propose test statistics that extract different features from the sample. Indeed, Theorem 3.1 states that if conditional correlation is different from zero even for only one quantile split, then the random variables are not independent. For example, if one wants to check the increase of dependence in the tails, then extreme tail splits might be considered, see Remark 3. We refer to [19] for a discussion on quantile split choices in the univariate setting and construction of test statistics based on sums/ratios of conditional coefficients.

Finally, we note that for simplicity we decided to focus on a bivariate setting but similar examples could be constructed for the multivariate setting using the characterisation presented in Theorem 3.2. Also, note that the main goal of the following examples is to show potential applications of the characterisation stated in Theorem 3.1. Consequently, for brevity, we decided to focus on intuitive illustrations rather than formal statistical methodology introduction. More systematic analysis of the listed application and development of the linked statistical frameworks is left for the future research.

Example 1 (Detecting non-linear dependence with conditional correlations)**.** We show an exemplary situation in which the conditional covariance (and correlations) coefficients could be explicitly computed and used to detect dependence structures even though the unconditional correlation is equal to zero. Let X be a standard normal random variable and let $Y := WX$, where W is independent of X and distributed uniformly on $\{-1, 1\}$. Then, (X, Y) is a vector with standard normal margins such that X and Y are uncorrelated but not independent (see [6]). In particular, the independence of X and Y cannot be (theoretically) rejected with the help of the classical Pearson correlation coefficient. However, Theorem 3.1 shows that the conditional corellations (or covariances) can be used to formally show the lack of independence. To illustrate this, let us provide an explicit formula for the quantile conditional covariances of X and Y. Fix the set A corresponding to the quantile splits $0 < p_1 < q_1 < 1$ and $0 < p_2 < q_2 < 1$, and define

$$
\begin{aligned}
l &:= \max(\Phi^{-1}(p_1), \Phi^{-1}(p_2))1_{\{q_1>p_2\}}1_{\{q_2>p_1\}},\\
r &:= \min(\Phi^{-1}(q_1), \Phi^{-1}(q_2))1_{\{q_1>p_2\}}1_{\{q_2>p_1\}},\\
\hat{l} &:= \max(\Phi^{-1}(p_1), \Phi^{-1}(1-q_2))1_{\{q_1>1-q_2\}}1_{\{1-p_2>p_1\}},\\
\hat{r} &:= \min(\Phi^{-1}(q_1), \Phi^{-1}(1-p_2))1_{\{q_1>1-q_2\}}1_{\{1-p_2>p_1\}};
\end{aligned}
$$

where Φ and ϕ denote cumulative distribution function and probability density function of standard normal random variable, respectively. Note that l and r are simply the left-most and the right-most points of the set $[\Phi^{-1}(p_1), \Phi^{-1}(q_1)] \cap [\Phi^{-1}(p_2), \Phi^{-1}(q_2)]$, respectively, provided that the intersection is non-empty; a

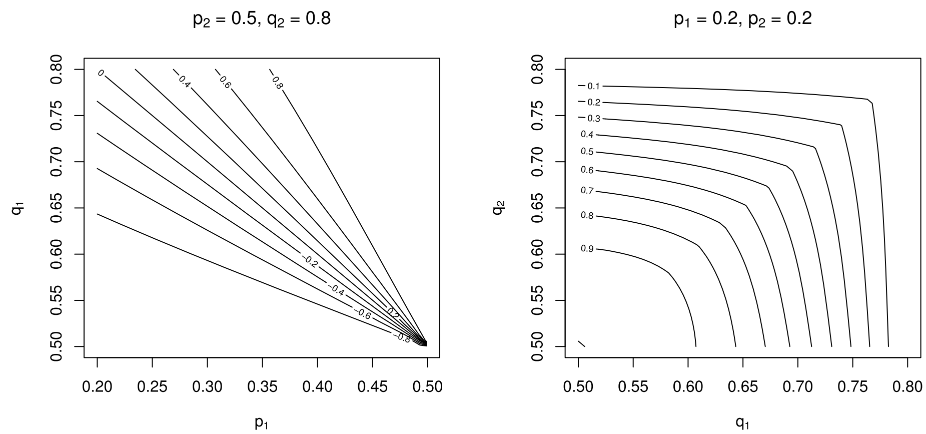


Fig 1: Contour plots of $\mathrm{Cor}_A[X,Y]$ for various choices of quantile splits based on Example 1 setting. Note that conditional covariance structure could be non-trivial even if $\mathrm{Cor}[X,Y]=0$.

similar interpretation holds for $\hat{l}$ and $\hat{r}$. Then, we get

$$\begin{aligned}
\mathrm{Cov}_A[X,Y] &= \frac{l\phi(l)-r\phi(r)+\Phi(r)-\Phi(l)}{\Phi(r)-\Phi(l)+\Phi(\hat{r})-\Phi(\hat{l})} - \frac{\hat{l}\phi(\hat{l})-\hat{r}\phi(\hat{r})+\Phi(\hat{r})-\Phi(\hat{l})}{\Phi(r)-\Phi(l)+\Phi(\hat{r})-\Phi(\hat{l})} \\
&\quad - \frac{(\phi(l)-\phi(r))^2-(\phi(\hat{l})-\phi(\hat{r}))^2}{(\Phi(r)-\Phi(l)+\Phi(\hat{r})-\Phi(\hat{l}))^2}, \\
\mathrm{Var}_A[X] &= \frac{l\phi(l)-r\phi(r)+\Phi(r)-\Phi(l)+\hat{l}\phi(\hat{l})-\hat{r}\phi(\hat{r})+\Phi(\hat{r})-\Phi(\hat{l})}{\Phi(r)-\Phi(l)+\Phi(\hat{r})-\Phi(\hat{l})} \\
&\quad - \frac{(\phi(l)-\phi(r)+\phi(\hat{l})-\phi(\hat{r}))^2}{(\Phi(r)-\Phi(l)+\Phi(\hat{r})-\Phi(\hat{l}))^2}, \\
\mathrm{Var}_A[Y] &= \frac{l\phi(l)-r\phi(r)+\Phi(r)-\Phi(l)+\hat{l}\phi(\hat{l})-\hat{r}\phi(\hat{r})+\Phi(\hat{r})-\Phi(\hat{l})}{\Phi(r)-\Phi(l)+\Phi(\hat{r})-\Phi(\hat{l})} \\
&\quad - \frac{(\phi(l)-\phi(r)-\phi(\hat{l})+\phi(\hat{r}))^2}{(\Phi(r)-\Phi(l)+\Phi(\hat{r})-\Phi(\hat{l}))^2}. \qquad (20)
\end{aligned}$$

In particular, for $p_1=p_2=0.5$ and $q_1=q_2=0.8$, we have $\mathrm{Cov}_A[X,Y]\approx 0.0573$ and $\mathrm{Cor}_A[X,Y]=1.0$, which directly proves that X and Y are not independent. For completeness, in Figure 1 we also present the values of $\mathrm{Cor}_A[X,Y]$ for exemplary quantile splits based on (20). From the figure we can see that the quantile conditional covariances may detect dependence between random variables, as described in Theorem 3.1.

Example 2 (Tests of statistical independence)**.** We present an exemplary statistical independence testing framework based on conditional correlations in a controlled environment. Let (X, Y) be a random vector with standard normal margins. Assume that under the null hypothesis X and Y are independent but under the alternative hypothesis the dependence structure is different. For testing purposes, we consider a test statistic (denoted by CondCor) that is given as the sample conditional correlation coefficient for the fixed quantile splits $p_1 = q_1 = 0.5$ and $p_2 = q_2 = 1.0$. We compare this test statistic with two alternative benchmarks. The first benchmark is the classical Pearson's correlation test statistic (denoted by Cor) given by the unconditional correlation coefficient. The second benchmark is the Hoeffding's independence test statistic (denoted by Hoeff) which checks if the empirical joint distribution function is a product of the marginal empirical distribution functions. For illustration, using Monte Carlo simulations, we check the test power for all three test statistics and various alternative hypotheses. Namely, let (N_1, N_2, W) be a random vector with mutually independent margins, where $N_1, N_2 \sim N(0, 1)$ and W is distributed uniformly on $\{-1, 1\}$. Given $a \in [0, 1]$, let

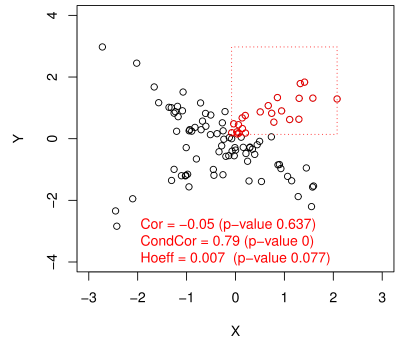
$$X := N_1, \quad \text{and} \quad Y := \sqrt{1-a}N_2 + \sqrt{a}WN_1. \tag{21}$$

We consider a null hypothesis $a = 0$ and alternative hypotheses $a = 0.25$, $a = 0.5$, $a = 0.75$, and $a = 1$. Note that the bigger the value of $a \in [0, 1]$, the stronger the (non-linear) dependence between X and Y, but we always get $\text{Cov}[X, Y] = 0$.

To confront exemplary test powers for all test statistics we consider sample size $n \in \{50, 100\}$ and the confidence level $\alpha = 0.05$. We perform Monte Carlo simulations of size 100 000 to simulate the null hypothesis distributions as well as to estimate the test power under different alternative hypotheses. More specifically, first, we simulate the null distribution of the respective test statistics based on the samples from independent standard normal random variables X and Y. Then, we compute the rejection thresholds and use them to estimate the power of the test by checking the proportion of the samples following the distribution of the random vector (X, Y) given by (21), for which we reject the null hypothesis. The results are given in Table 2. The presented results suggest a very good performance of the proposed methodology, especially when a is close to 1. To better illustrate these results, in Figure 2 we also present an exemplary dataset when only the CondCor test rejected the null independence hypothesis. As seen in the figure, the conditional correlation approach should be effective for regions in which some form of dependence is visible.

Example 3 (Detecting serial dependence)**.** In this example, we show how to use conditional correlations to detect serial dependence in time series. Note that the analysis of (unconditional) auto-correlation is a standard technique used in time series analysis and signal processing to detect serial dependence, see e.g. [14, 5] and references therein. In particular, it is often used to verify the lack of trend or volatility clustering in financial data, see e.g. [11, 8, 20]. Given a time-series sample $(x_t)_{t=1}^n$, the empirical auto-correlation of lag k is typically computed by estimating the (unconditional) correlation between the sub-samples $(x_t)_{t=k+1}^n$ and $(x_t)_{t=1}^{n-k}$. In this simple example, we use market data to show how

n	H_1	Cor	Hoeff	CondCor
50	$a = 0.25$	10.5%	6.8%	**10.7%**
	$a = 0.50$	16.1%	10.3%	**29.2%**
	$a = 0.75$	21.4%	18.5%	**70.4%**
	$a = 1.00$	25.8%	**100%**	**100%**
100	$a = 0.25$	10.9%	7.2%	**18.5%**
	$a = 0.50$	16.5%	12.6%	**58.3%**
	$a = 0.75$	21.5%	33.3%	**96.4%**
	$a = 1.00$	25.7%	**100%**	**100%**



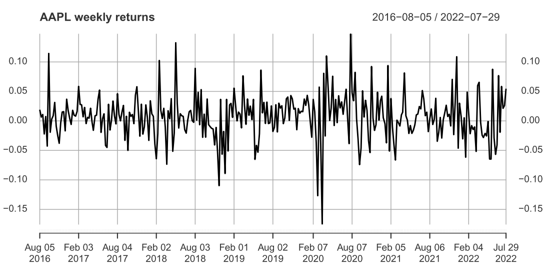
(a) (b)

Fig 2: The table (a) presents test powers for various choices of $a \in [0, 1]$ under the settings described in Example 2. The best performance is marked in bold – one can see that in all cases CondCor outperforms other frameworks. The plot (b) presents exemplary sample for $n = 100$ and $a = 0.75$ on which only CondCor test statistic rejected the null hypothesis. The red region indicates the subsample on which empirical conditional correlation is computed.

the information about conditional correlation could be used to refine standard auto-correlation function (ACF) analysis. For simplicity, we decided to take one exemplary stock market data. Namely, we consider weekly (adjusted) log-returns of AAPL stock prices in the period 01/08/2016 – 01/08/2022, which we denote by $(r_t)_{t=1}^n$, $n = 312$, the data is illustrated in Figure 3. In Figure 4, we present the classical auto-correlation function (ACF) plots for log-returns, absolute values of log-returns, as well as squared log-returns. While, the first plot is often used for generic independence check (lack of trend), the last two might be used to investigate the so-called volatility clustering effect, see [8]. Although from Figure 3 one can deduce that the data is not i.i.d. (for lag $k = 3$ the 1% i.i.d. confidence threshold level is breached), the ACF functions do not detect any major problem for lag $k = 1$, suggesting no obvious dependence between consequent observations. Let us now focus on lag $k = 1$ and check if we can refine the ACF analysis using conditional equivalent of auto-correlation. Namely, let us consider the quantile split $p_1 = p_2 = 0.01$ and $q_1 = q_2 = 0.7$, and compute the conditional correlation on the corresponding set A given by (3), for the samples $(r_t)_{t=2}^n$ and $(r_t)_{t=1}^{n-1}$, see Figure 5 for illustration. Note that this particular quantile split could be linked to a potential dependence structure visible in the scatter plot – this is an exemplary region, where data seems to be positively correlated.

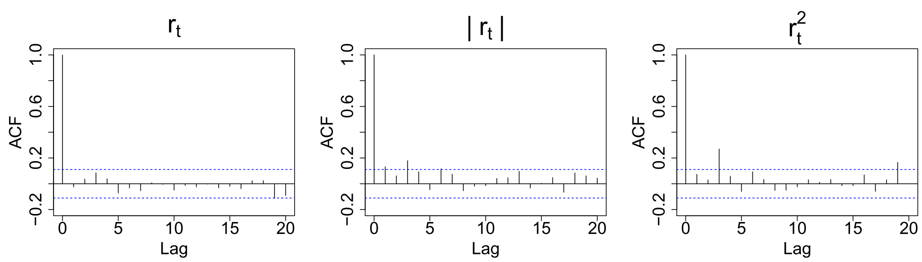
The estimated value of the conditional correlation is equal to 0.31, which indicates that the time-series observations are not independent. To sanity check if this claim is statistically significant, we performed a simple normal distribution-based Monte Carlo exercise. Namely, we picked $M = 100\,000$ strong Monte Carlo

(a)



(b)

Fig 3: AAPL stock price daily quotes (a) and weekly log-returns (b) in the period 01/08/2016 – 01/08/202 used in Example 3.



(a) (b) (c)

Fig 4: AAPL stock price weekly log-returns in the period 01/08/2016 – 01/08/2022. Auto-correlation function (ACF) plots for (a) log-returns, (b) absolute values of log-returns, (c) squared log-returns.

samples of size $n = 312$, i.e. size equal to the size of the original sample, from independent normal distributions. For each run, we computed the conditional correlation for the same lag and the same sample quantile set. The 0.1% upper quantile of the obtained MC density is equal to 0.25, which shows that the initial sample empirical conditional correlation 0.31 is (statistically) significantly different from zero which proved serial dependence between consequent observations.

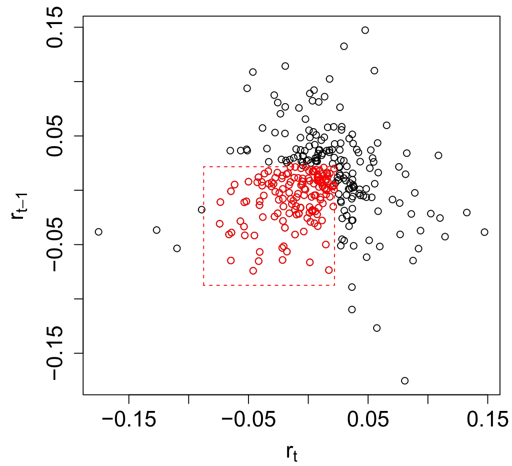


Fig 5: Lag plot ($k = 1$) for weekly APPL log-returns. The red region indicates the conditional set A on which we computed (empirical) conditional correlation which is equal to 0.31. This value is statistically different from zero which indicates serial dependence of consequent observations.

Example 4 (Testing independence in regression models)**.** The results from this paper could be used to develop statistical tests for residual dependence structure checks in the regression models. For example, in the linear regression, one often has to check if the explanatory variables are independent of errors which values are proxied by residuals. The usage of conditional correlations facilitates detection and formal testing dependence structures that are non-linear. To illustrate this, let us take the Nerlove's 1955 dataset with cost function for electricity producers as the output variable, see [24]. Our goal is to check the dependence between the residuals and the output for the Cobb-Douglas model fitted to this dataset, see pages 76–84 in [15] for exact model construction and details. We consider the quantile split $p_1 = p_2 = 0.1$ and $q_1 = q_2 = 0.9$, which should provide a good balance between the sample size loss and the possible outliers exclusion. Then, we compute the conditional correlation on the corresponding set A for the vector of residuals and logged output, see Figure 6 for data illustration. While the unconditional correlation for this data is equal to zero (by model construction), the empirical conditional correlation is equal to 0.45 which indicates non-linear dependence between the output and residuals as pointed out in [15], where this conclusion was reached through visual inspection. Statistical significance of non-zero correlation could be checked using a similar logic as in Example 3.

Example 5 (Conditional and partial independence)**.** As already noted in Remark 4, conditional correlations could be used to characterise conditional independence as well as partial independence; see e.g [4] for further discussion

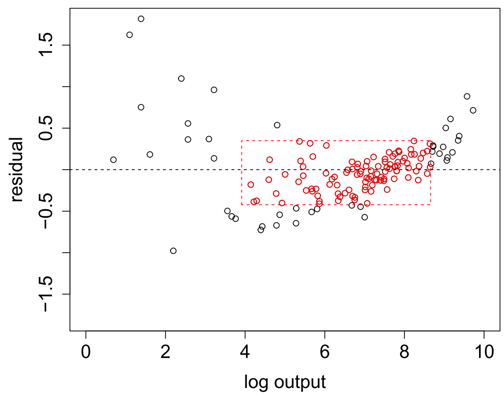


Fig 6: The plot presents the residuals plotted against the log output for the Cobb-Douglas model applied to Marlov's dataset, see Figure 1.7 in [15] for reference. While the unconditional empirical correlation is equal to zero, the conditional empirical correlation in the red region is equal to 0.45 which shows that residuals are not independent from the model output.

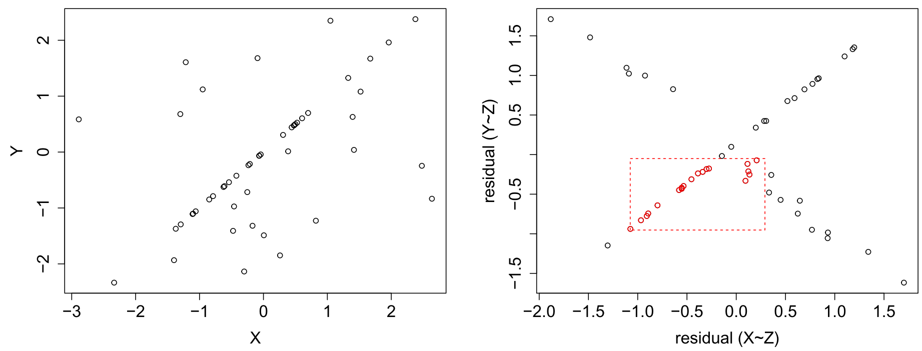
on these concepts. In this example, we provide a toy example in which partial correlation is equal to zero but non-linear dependence is visible. Following a logic similar to the one presented in Example 2 let us consider a random vector (X, Y, Z) given by

$$X := N_1 + N_2, \quad Y := WN_1 + N_2, \quad \text{and} \quad Z := N_2,$$

where (N_1, N_2, W) is a random vector with mutually independent margins, $N_1, N_2 \sim N(0, 1)$ and W is distributed uniformly on $\{-1, 1\}$. We treat Z as the control variable. It is easy to check that $\mathrm{Cor}(X, Y) = 0.5$ due to the effect of Z on X and Y. On the other hand, the analysis of partial correlation of X and Y, given Z, shows partial linear independence. Indeed, we have

$$\rho_{XY\cdot Z} = \frac{\mathrm{Cor}[X, Y] - \mathrm{Cor}[X, Z]\,\mathrm{Cor}[Y, Z]}{\sqrt{1 - \mathrm{Cor}[X, Z]^2}\sqrt{1 - \mathrm{Cor}[Y, Z]^2}} = \frac{0.5 - \sqrt{0.5}\sqrt{0.5}}{\sqrt{0.5}\sqrt{0.5}} = 0,$$

but clearly X and Y are non-linearly dependent even when taking into account information from Z. To understand this, it is enough to recall that $\rho_{XY\cdot Z}$ measures the correlation of residual errors from two linear regressions models, with X and Y as response and Z as regressor. Noting that those errors form a bivariate vector (N_1, WN_1), we see that one can detect non-linear dependence between those residuals using conditional correlation analysis as done in Example 1. To illustrate this we picked a random sample from (X, Y, Z) of size $n = 50$. The sample empirical correlation between X and Y is equal to 0.47, while the empirical partial correlation is equal to -0.08 which is consistent with theoretical values, see Figure 7 for data illustration. We also estimated conditional



(a) (b)

Fig 7: Exemplary sample from (X, Y) of size $n = 50$ under the setting from Example 5 is presented in (a). The empirical correlation between X and Y is equal to 0.47 since both X and Y depend on control variable Z. The residuals from regresing the margins on Z are presented in (b). While the empirical partial correlation is equal to -0.08 as X and Y and partially linearly independent, the conditional partial correlation on the red set is equal to 0.89.

(partial) correlation for splits $p_1 = p_2 = 0.1$ and $q_1 = q_2 = 0.5$ and obtained the value 0.89. As expected, this indicates a strong non-linear (local) dependence between X and Y, even when the effect of Z is removed.

5. Concluding remarks

In this paper we provided an alternative characterisation of statistical independence based on conditional correlations in both bivariate and multivariate setting, see Theorem 3.1 and Theorem 3.2. We showed that any form of dependence between random variables could be described using a family of localised Pearson's correlations, i.e. non-linear dependence could be described by local linear dependence via conditional correlations. It is worth noting that the idea of localised correlation is relatively simple, e.g. when confronted with higher-order moments based measures of association or advanced copula-based tools, so that this characterisation could be appealing to engineers which are used to linear dependence measurement. We also showed that the characterisation based on conditional correlations could be potentially used to refine multiple existing statistical frameworks in which independence is analysed. Our claim was illustrated using numerous toy examples, see Section 4. We want to emphasize that this work is a first step towards developing formal statistical methodologies for evaluating independence based on local correlations. As such, it has potential to open new theoretical and practical research areas.

Funding

Marcin Pitera acknowledges support from the National Science Centre, Poland, via project 2020/37/B/HS4/00120. Part of the work of Damian Jelito was funded by the Priority Research Area Digiworld under the program Excellence Initiative – Research University at the Jagiellonian University in Kraków.

References

[1] Aielli, G. P. (2013). Dynamic conditional correlation: on properties and estimation. *Journal of Business & Economic Statistics* **31** 282–299.

[2] Akemann, C. A., Bruckner, A. M., Robertson, J. B., Simons, S. and Weiss, M. L. (1984). Asymptotic conditional correlation coefficients for truncated data. *Journal of Mathematical Analysis and Applications* **99** 350–434.

[3] Aldrich, J. (1995). Correlations genuine and spurious in Pearson and Yule. *Statistical Science* 364–376.

[4] Baba, K., Shibata, R. and Sibuya, M. (2004). Partial correlation and conditional correlation as measures of conditional independence. *Australian & New Zealand Journal of Statistics* **46** 657–664.

[5] Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting*, 3rd ed. Springer.

[6] Broffitt, J. D. (1986). Zero correlation, independence, and normality. *The American Statistician* **40** 276–277.

[7] Chernoff, P. R., Mori, T. F., Szanto, S., Erugin, N. P. and Evans, R. J. (1981). Advanced Problems: 6326-6329. *The American Mathematical Monthly* **88** 68–69.

[8] Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance* **1** 223.

[9] David, H. A. (2009). A historical note on zero correlation and independence. *The American Statistician* **63** 185–186.

[10] Durante, F. and Jaworski, P. (2010). Spatial contagion between financial markets: a copula-based approach. *Applied Stochastic Models in Business and Industry* **26** 551–564.

[11] Fama, E. F. and French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy* **96** 246–273.

[12] Galton, F. (1889). *Natural Inheritance*. Macmillan and Company.

[13] Gretton, A., Smola, A., Bousquet, O., Herbrich, R., Belitski, A., Augath, M., Murayama, Y., Pauls, J., Schölkopf, B. and Logothetis, N. (2005). Kernel Constrained Covariance for Dependence Measurement. In *Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics* (R. G. Cowell and Z. Ghahramani, eds.). *Proceedings of Machine Learning Research* **R5** 112–119. PMLR.

[14] Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

[15] Hayashi, F. (2000). *Econometrics*. Princeton University Press.

[16] Hebda-Sobkowicz, J., Zimroz, R., Pitera, M. and Wyłomańska, A. (2020). Informative frequency band selection in the presence of non-Gaussian noise – a novel approach based on the conditional variance statistic with application to bearing fault diagnosis. *Mechanical Systems and Signal Processing* **145** 106971.

[17] Jaworski, P. and Pitera, M. (2016). The 20-60-20 Rule. *Discrete & Continuous Dynamical Systems-Series B* **21**.

[18] Jaworski, P. and Pitera, M. (2020). A note on conditional variance and characterization of probability distributions. *Statistics & Probability Letters* **163** 108800.

[19] Jelito, D. and Pitera, M. (2021). New fat-tail normality test based on conditional second moments with applications to finance. *Statistical Papers* **62** 2083–2108.

[20] Jiang, H., Saart, P. W. and Xia, Y. (2016). Asymmetric conditional correlations in stock returns. *The Annals of Applied Statistics* **10** 989 – 1018.

[21] Kenett, D. Y., Huang, X., Vodenska, I., Havlin, S. and Stanley, H. E. (2015). Partial correlation analysis: Applications for financial markets. *Quantitative Finance* **15** 569–578.

[22] Kotz, S. and Drouet, D. (2001). *Correlation and Dependence*. World Scientific.

[23] Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer New York.

[24] Nerlove, M. (1963). Returns to Scale in Electricity Supply. In "Measurement in Economics-Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld", edited by C. F. Christ. *Stanford University Press*.

[25] Pitera, M., Chechkin, A. and Wyłomanska, A. (2022). Goodness-of-fit test for α-stable distribution based on the quantile conditional variance statistics. *Statistical Methods & Applications* **31** 387–424.

[26] Rabinowitz, S. (1992). *Index to Mathematical Problems, 1980–1984* **1**. MathPro Press.

[27] Rao, M., Seth, S., Xu, J., Chen, Y., Tagare, H. and Príncipe, J. C. (2011). A test of independence based on a generalized correlation function. *Signal Processing* **91** 15–27.

[28] Scarsini, M. (1984). On measures of concordance. *Stochastica* **8** 201–218.

[29] Shiryaev, A. N. (1996). *Probability*. Springer.

[30] Stigler, S. M. (1989). Francis Galton's account of the invention of correlation. *Statistical Science* 73–79.

[31] Székely, G. J., Rizzo, M. L. and Bakirov, N. K. (2007). Measuring and testing dependence by correlation of distances. *The Annals of Statistics* **35** 2769–2794.

[32] Tjøstheim, D. and Hufthammer, K. O. (2013). Local Gaussian correlation: A new measure of dependence. *Journal of Econometrics* **172** 33–48.

[33] Tjøstheim, D., Otneim, H. and Støve, B. (2022). Statistical Dependence: Beyond Pearson's ρ. *Statistical Science* **37** 90–109.

[34] Veličković, V. (2015). What everyone should know about statistical correlation. *American Scientist* **103** 26–29.
[35] Witsenhausen, H. S. (1975). On sequences of pairs of dependent random variables. *SIAM Journal on Applied Mathematics* **28** 100–113.
[36] Zhu, L., Xu, K., Li, R. and Zhong, W. (2017). Projection correlation between two random vectors. *Biometrika* **104** 829–843.