Measuring monotony in two-dimensional samples

Farida Kachapova[a*] and Ilias Kachapov[b]

[a] *School of Computing and Mathematical Sciences, Auckland University of Technology, Auckland, New Zealand;* [b] *University of Auckland, Auckland, New Zealand*

* Corresponding author. Email: farida.kachapova@aut.ac.nz

The paper introduces a monotony coefficient as a new measure of the monotone dependence in a two-dimensional sample. Some properties of this measure are derived. In particular, it is shown that the absolute value of the monotony coefficient for a two-dimensional sample is between $|r|$ and 1, where r is the Pearson's correlation coefficient for the sample; that the monotony coefficient equals 1 for any monotone increasing sample and equals -1 for any monotone decreasing sample. The paper contains a few examples demonstrating that the monotony coefficient is a more accurate measure of the degree of monotone dependence for a non-linear relationship than the Pearson's, Spearman's and Kendall's correlation coefficients. The monotony coefficient is a tool that can be applied to samples in order to find dependencies between random variables; it is especially useful in finding couples of dependent variables in a big dataset of many variables. Undergraduate students in mathematics and science would benefit from learning and applying this measure of monotone dependence.

Keywords: monotony; dependence; correlation

1. Introduction

There are several measures of dependence between two random variables. The best known is the Pearson's correlation coefficient ρ. It measures the degree of linear dependence between two random variables. Linear models make quite a narrow class. Students in mathematics and science would benefit from studying other types of dependence, e.g. monotone dependence.

Gebelein [1] introduced a measure ρ' of dependence between two non-degenerate random variables X and Y by the following:

$$\rho'(X, Y) = \sup \rho\,[\,f(X), g(Y)],$$

where the supremum is taken over all Borel-measurable functions f, g, for which

$$0 < \operatorname{Var} f(X) < \infty \quad \text{and} \quad 0 < \operatorname{Var} g(Y) < \infty.$$

Kimeldorf and Sampson [2] introduced the monotone correlation ρ^* between two non-degenerate random variables X and Y by the following:

$$\rho^*(X, Y) = \sup \rho\,[\,f(X), g(Y)],$$

where the supremum is taken over all monotone functions f, g, for which

$$0 < \operatorname{Var} f(X) < \infty \quad \text{and} \quad 0 < \operatorname{Var} g(Y) < \infty.$$

ρ^* is a more useful measure of dependence than ρ'. In particular, Kimeldorf and Sampson [2] showed that for non-degenerate random variables X and Y,

$\rho^*(X, Y) = 0$ if and only if X and Y are independent. They also showed that if X and Y are monotone dependent, then $\rho^*(X, Y) = 1$; but the converse implication is false.

Mayer [3] introduced another monotone correlation coefficient m_{xy} (defined in terms of a supremum over all monotone functions) and used it to estimate ρ.

Thus, the existing monotone correlation coefficients are defined for random variables in terms of a supremum, that is in non-constructive (non-algorithmic) way. Such coefficients cannot be easily calculated and applied.

For samples the most common measures of dependence are the Pearson's sample correlation coefficient, Spearman's and Kendall's rank correlation coefficients (see, for example, [4]). The last two statistics are used for ordinal data and numerical data converted to rankings. These two statistics measure the degree of monotone dependence between the ranks of the data. We did not find in literature a more accurate sample measure of monotony for continuous data.

In this paper we introduce the monotony coefficient rm, which measures the degree of monotone dependence in a finite two-dimensional sample without converting it to rankings. The monotone dependence measured by this coefficient is a more general and important relation than the linear dependence measured by the Pearson correlation. The monotony coefficient is defined in simple terms and the proof of its properties involves only simple algebra. So the monotony coefficient can be included in undergraduate statistics courses along with the Pearson correlation. It would expand the students' perspective on statistical dependencies.

2. Notations

For a sample $x = (x_1, x_2,\ldots, x_n)$ we will use notations $\bar{x}$ for the sample mean and s_x for the sample standard deviation. For a two-dimensional sample $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix},\ldots,\begin{pmatrix} x_n \\ y_n \end{pmatrix}$ there are two corresponding one-dimensional samples: $x = (x_1,\ldots, x_n)$ and $y = (y_1,\ldots, y_n)$, so in short we denote the two-dimensional sample as $\begin{pmatrix} x \\ y \end{pmatrix}$. The sample covariance is defined by $s(x, y) = \dfrac{1}{n-1}\sum\limits_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$. The Pearson's coefficient of correlation is defined by $r(x, y) = \dfrac{s(x, y)}{s_x \cdot s_y}$.

The sample $x = (x_1, x_2,\ldots, x_n)$ is called **constant** if $x_1 = x_2 = \ldots = x_n$.

3. Monotone covariance

First we introduce a new type of sample covariance sm (where m stands for monotone).

Definition 1. Suppose $x = (x_1, x_2,\ldots, x_n)$ is a finite sample.

1) $x^* = (x_1^*, x_2^*,\ldots, x_n^*)$ denotes the sample of the same numbers in ascending order: $x_1^* \leq x_2^* \leq \ldots \leq x_n^*$.

2) $x' = (x_1', x_2',\ldots, x_n')$ denotes the sample of the same numbers in descending order: $x_1' \geq x_2' \geq \ldots \geq x_n'$.　　　　　　　　　•

Definition 2. Suppose $\begin{pmatrix} x \\ y \end{pmatrix}$ is a two-dimensional sample. Its **monotone covariance**

sm is defined by: $\quad sm\ (x, y) = \begin{cases} s(x^*, y^*)\ if\ s(x, y) > 0, \\ \quad 0\ if\ s(x, y) = 0, \\ s(x^*, y')\ if\ s(x, y) < 0. \end{cases}$ ●

Definition 3. Suppose $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, ..., \begin{pmatrix} x_n \\ y_n \end{pmatrix}$ is a two-dimensional sample.

    1) The sample is called **monotone increasing** if both samples x and y are not constant and for any $i, j = 1, 2, ..., n$,

$$x_i < x_j \quad \Rightarrow \quad y_i \le y_j.$$

    2) The sample is called **monotone decreasing** if both samples x and y are not constant and for any $i, j = 1, 2, ..., n$,

$$x_i < x_j \quad \Rightarrow \quad y_i \ge y_j.$$

    3) The sample is called **monotone** if it is monotone increasing or decreasing. ●
Obviously the Definition 3 is symmetrical with respect to x and y.
To study properties of the monotone covariance, we need the following lemmas.

Lemma 1. $s\ (x^*, y^*) = s\ (x', y\ ').$

Lemma 2. $s\ (x^*, y\ ') = s\ (x', y^*).$

Lemma 3. Suppose $\begin{pmatrix} u \\ v \end{pmatrix}$ is a two-dimensional sample.

    1) If $\begin{pmatrix} u \\ v \end{pmatrix}$ is monotone increasing, $\sum\limits_{i=1}^{n} u_i = 0$ and $\sum\limits_{i=1}^{n} v_i = 0$, then $\sum\limits_{i=1}^{n} u_i\, v_i > 0.$

    2) If $\begin{pmatrix} u \\ v \end{pmatrix}$ is monotone decreasing, $\sum\limits_{i=1}^{n} u_i = 0$ and $\sum\limits_{i=1}^{n} v_i = 0$, then $\sum\limits_{i=1}^{n} u_i\, v_i < 0.$

Lemma 4. Suppose $\begin{pmatrix} x \\ y \end{pmatrix}$ is a two-dimensional sample.

    1) If $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone increasing, then $s\ (x, y) > 0.$

    2) If $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone decreasing, then $s\ (x, y) < 0.$

Lemma 5. If the samples x and y are not constant, then

$$s\ (x^*, y^*) > 0 \ \text{ and } \ s\ (x^*, y\ ') < 0.$$

Lemma 6. 1) $s(x, y) \le s(x^*, y^*)$.

2) $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone increasing $\Leftrightarrow$ $(s(x, y) > 0$ and $s(x, y) = s(x^*, y^*))$.

Lemma 7. 1) $s(x^*, y') \le s(x, y)$.

2) $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone decreasing $\Leftrightarrow$ $(s(x, y) < 0$ and $s(x^*, y') = s(x, y))$.

The proofs for these lemmas are given in [5]. They involve only basic algebra and mathematical induction; each lemma follows from the previous ones.

Theorem 1. Properties of the monotone covariance. For the two-dimensional

sample $\begin{pmatrix} x \\ y \end{pmatrix}$ the following holds.

1) $|s(x, y)| \le |sm(x, y)| \le s_x \cdot s_y$.

2) $sm(x, y) = 0 \quad \Leftrightarrow \quad s(x, y) = 0$.

3) $sm(x, y) > 0 \quad \Leftrightarrow \quad s(x, y) > 0$.

4) $sm(x, y) < 0 \quad \Leftrightarrow \quad s(x, y) < 0$.

5) $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone increasing $\Leftrightarrow$ $(s(x, y) > 0$ and $sm(x, y) = s(x, y))$.

6) $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone decreasing $\Leftrightarrow$ $(s(x, y) < 0$ and $sm(x, y) = s(x, y))$.

7) $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone $\Leftrightarrow$ $(s(x, y) \ne 0$ and $sm(x, y) = s(x, y))$.

Proof: 1) First we prove the left inequality. If $s(x, y) = 0$, then it is obvious. Assume $s(x, y) \ne 0$. If $s(x, y) > 0$, then by Lemma 6.1), $0 < s(x, y) \le s(x^*, y^*) = sm(x, y)$. If $s(x, y) < 0$, then $sm(x, y) = s(x^*, y') \le s(x, y) < 0$ by Lemma 7.1). In both cases $|s(x, y)| \le |sm(x, y)|$.

Now we will prove the right inequality. Obviously, $s_{x^*} = s_x$, $s_{y^*} = s_{y'} = s_y$. By a well-known property of the covariance, $|s(x^*, y^*)| \le s_x \cdot s_y$ and $|s(x^*, y')| \le s_x \cdot s_y$, which proves $|sm(x, y)| \le s_x \cdot s_y$.

2) $\Rightarrow$ Suppose $sm(x, y) = 0$. By 1), $|s(x, y)| \le |sm(x, y)| = 0$, so $s(x, y) = 0$.

$\Leftarrow$ Suppose $s(x, y) = 0$. Then by the definition $sm(x, y) = 0$.

3) $\Rightarrow$ Suppose $sm(x, y) > 0$. Then by 2), $s(x, y) \ne 0$ and the samples x and y are

not constant. If $s(x, y) < 0$, then $sm(x, y) = s(x^*, y') < 0$ by Lemma 5; contradiction. So $s(x, y) > 0$.

$\Leftarrow$ Suppose $s(x, y) > 0$. Then the samples x and y are not constant and $sm(x, y) = s(x^*, y^*) > 0$ by Lemma 5.

4) follows immediately from 2) and 3).

5) $\Rightarrow$ Suppose $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone increasing. Then by Lemma 6.2), $s(x, y) > 0$ and $sm(x, y) = s(x^*, y^*) = s(x, y)$.

$\Leftarrow$ Suppose $s(x, y) > 0$ and $sm(x, y) = s(x, y)$. Then $s(x^*, y^*) = s(x, y)$ and $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone increasing by Lemma 6.2).

6) $\Rightarrow$ Suppose $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone decreasing. Then by Lemma 7.2), $s(x, y) < 0$ and $sm(x, y) = s(x^*, y') = s(x, y)$.

$\Leftarrow$ Suppose $s(x, y) < 0$ and $sm(x, y) = s(x, y)$. Then $s(x^*, y') = s(x, y)$ and $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone decreasing by Lemma 7.2).

7) $\Rightarrow$ Suppose $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone. Then it is monotone increasing or monotone decreasing. In both cases $s(x, y) \neq 0$ and $sm(x, y) = s(x, y)$ by 5) and 6).

$\Leftarrow$ Suppose $s(x, y) \neq 0$ and $sm(x, y) = s(x, y)$. If $s(x, y) > 0$, then $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone by 5). If $s(x, y) < 0$, then $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone by 6). ∎

Part 5) of Theorem 1 says, in particular, that $s(x, y) > 0$ for any monotone increasing sample $\begin{pmatrix} x \\ y \end{pmatrix}$. It is a well-known property of the covariance, but it is hard to find proof for it in literature. A similar, but a non-strict inequality is proven for the population covariance in [6] and [7].

4. Monotony coefficient

Next we introduce a normalized measure of monotony.

Definition 4. The **monotony coefficient** *rm* is defined by:

$$rm(x, y) = \begin{cases} 0 & \text{if } s(x, y) = 0, \\ \dfrac{|s(x, y)|}{sm(x, y)} & \text{if } s(x, y) \neq 0. \end{cases}$$ •

The definition is valid due to Theorem 1.2).

When the coefficient $rm(x, y)$ is closer to 1, then the sample covariance $s(x, y)$ is closer to the covariance $s(x^*, y^*)$ of an increasing sample, hence the sample $\begin{pmatrix} x \\ y \end{pmatrix}$ is more monotone increasing. Similarly, when the coefficient $rm(x, y)$ is closer to -1, the sample covariance $s(x, y)$ is closer to the covariance $s(x^*, y')$ of a decreasing sample, hence the sample $\begin{pmatrix} x \\ y \end{pmatrix}$ is more monotone decreasing.

Theorem 2. Properties of the monotony coefficient. For the two-dimensional sample $\begin{pmatrix} x \\ y \end{pmatrix}$ the following holds.

   1) $|r(x, y)| \leq |rm(x, y)| \leq 1$ when $r(x, y)$ is defined.

   2) $rm(x, y) = 1$ if and only if the sample is monotone increasing.

   3) $rm(x, y) = -1$ if and only if the sample is monotone decreasing.

Proof: 1) If $s(x, y) = 0$, then $sm(x, y) = 0$ and $r(x, y) = rm(x, y) = 0$.

Assume $s(x, y) \neq 0$. By Theorem 1.1), $0 < |s(x, y)| \leq |sm(x, y)| \leq s_x \cdot s_y$.

So $\dfrac{1}{s_x \cdot s_y} \leq \dfrac{1}{|sm(x, y)|} \leq \dfrac{1}{|s(x, y)|}$.

Multiply all sides of this inequality by $|s(x, y)|$: $\dfrac{|s(x, y)|}{s_x \cdot s_y} \leq \dfrac{|s(x, y)|}{|sm(x, y)|} \leq 1$, which gives $|r(x, y)| \leq |rm(x, y)| \leq 1$.

2) $\Rightarrow$ Suppose $rm(x, y) = 1$. Then $s(x, y) \neq 0$, $sm(x, y) = |s(x, y)|$, so $sm(x, y) > 0$ and by Theorem 1.3), $s(x, y) > 0$. Therefore $sm(x, y) = s(x, y)$ and by Theorem 1.5), $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone increasing.

$\Leftarrow$ Suppose $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone increasing. Then by Theorem 1.5), $s(x, y) > 0$ and $sm(x, y) = s(x, y)$. So $rm(x, y) = \dfrac{|s(x, y)|}{sm(x, y)} = \dfrac{s(x, y)}{s(x, y)} = 1$.

3) $\Rightarrow$ Suppose $rm(x, y) = -1$. Then $s(x, y) \neq 0$, $sm(x, y) = -|s(x, y)|$, so $sm(x, y) < 0$ and by Theorem 1.4), $s(x, y) < 0$. Therefore $sm(x, y) = s(x, y)$ and by Theorem 1.6), $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone decreasing.

$\Leftarrow$ Suppose $\begin{pmatrix} x \\ y \end{pmatrix}$ is monotone decreasing. Then by Theorem 1.6), $s\,(x, y) < 0$

and $sm\,(x, y) = s\,(x, y)$. So $rm\,(x, y) = \dfrac{|s(x, y)|}{sm(x, y)} = \dfrac{-s(x, y)}{s(x, y)} = -1$. ∎

Example 1. We will calculate the monotony coefficients for the following two-dimensional samples.

a) $\begin{pmatrix} -1 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}$.

$s\,(x, y) = 0.5 > 0$. $\quad x^* = (-1, -1, 0, 0, 2)$, $\quad y^* = (-2, 0, 0, 1, 1)$, $\quad sm\,(x, y) = s\,(x^*, y^*) = 1$.

So $rm\,(x, y) = \dfrac{|s(x, y)|}{sm(x, y)} = \dfrac{0.5}{1} = 0.5$. Compare this with $r\,(x, y) \approx 0.3333$.

b) $\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

$s\,(x, y) = -0.4 < 0$. $\quad x^* = (-1, -1, 0, 0, 1, 1)$, $\quad y' = (1, 1, 1, -1, -1, -1)$,

$sm\,(x, y) = s\,(x^*, y') = -0.8$.

So $rm\,(x, y) = \dfrac{|s(x, y)|}{sm(x, y)} = \dfrac{0.4}{-0.8} = -0.5$. Compare this with $r\,(x, y) \approx -0.4082$.

The sample a) is positively correlated, hence mostly increasing, and the sample b) is negatively correlated, hence mostly decreasing. ∎

Example 2. The sample $\begin{pmatrix} 0.01 \\ 100 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 100 \\ 0.01 \end{pmatrix}$ is monotone decreasing. So

$rm\,(x, y) = -1$. But $r\,(x, y) \approx -0.35099$ and the coefficient of determination $r^2\,(x, y) \approx$ $\approx 0.12 = 12\%$. The traditional analysis of these data that uses the Pearson's coefficient will not recognize the dependence between the two variables, since the value of 0.35099 is small. But the monotony coefficient reflects the dependence; this dependence can be modelled by the function $y = \dfrac{1}{x}$. The relationship becomes linear after a logarithmic transformation but it is a non-trivial task to find a suitable transformation. ∎

Example 3. Consider two samples: $\begin{pmatrix} 0.5 \\ -1 \end{pmatrix}, \begin{pmatrix} 0.8 \\ -0.01 \end{pmatrix}, \begin{pmatrix} 1.1 \\ -0.05 \end{pmatrix}, \begin{pmatrix} 1.4 \\ 0.35 \end{pmatrix}, \begin{pmatrix} 1.7 \\ 0.34 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

and $\begin{pmatrix} 0.5 \\ -1 \end{pmatrix}, \begin{pmatrix} 0.8 \\ 0.199 \end{pmatrix}, \begin{pmatrix} 1.1 \\ -0.9 \end{pmatrix}, \begin{pmatrix} 1.4 \\ 0.9 \end{pmatrix}, \begin{pmatrix} 1.7 \\ 0.2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. These are the corresponding graphs:

Figure.

Neither of the samples is monotone but intuitively Sample 1 has more monotony than Sample 2. After converting both x and y values to ranks we get the same table for both samples:

Table.

The Spearman's correlation coefficient for each sample equals $(31)/(35) \approx 0.886$. The Kendall's correlation coefficient for each sample equals $(11)/(15) \approx 0.733$. So these coefficients do not reflect the difference in the monotony of the samples. But the monotony coefficient rm does: for the first sample rm is approximately 0.991 and for the second sample rm is approximately 0.766. Both samples are mostly increasing but the monotony coefficient of the first sample is closer to 1 showing that the first sample is more monotone. ∎

The relation between the monotony and Pearson coefficients is described in Theorem 3 below.

Definition 5. Suppose $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, ..., \begin{pmatrix} x_n \\ y_n \end{pmatrix}$ is a two-dimensional sample.

The samples x and y are said to have a **linear relation** if for some numbers a and b, the following holds:

$y_i = a + b\,x_i$ for all $i = 1, 2,\ldots, n$ or $x_i = a + b\,y_i$ for all $i = 1, 2,\ldots, n$.

This linear relation is called **positive** if $b > 0$ and **negative** if $b < 0$. •

Theorem 3. Relation to the Pearson correlation. For the two-dimensional sample $\begin{pmatrix} x \\ y \end{pmatrix}$ the following holds.

1) If $s\,(x, y) > 0$, then

$$rm\,(x, y) = r\,(x, y) \iff x^* \text{ and } y^* \text{ have a positive linear relation.}$$

2) If $s\,(x, y) < 0$, then

$$rm\,(x, y) = r\,(x, y) \iff x^* \text{ and } y' \text{ have a negative linear relation.}$$

3) If $s\,(x, y) = 0$, then $rm\,(x, y) = r\,(x, y) = 0$.

Proof: 1) Suppose $s\,(x, y) > 0$. Then $rm\,(x, y) = \dfrac{|s(x, y)|}{sm(x, y)} = \dfrac{s(x, y)}{s(x^*, y^*)}$. So

$$rm\,(x, y) = r\,(x, y) \iff \frac{s(x, y)}{s(x^*, y^*)} = \frac{s(x, y)}{s_x \cdot s_y} \iff s\,(x^*, y^*) = s_x \cdot s_y \iff s\,(x^*, y^*) = s_{x^*} \cdot s_{y^*}$$

$$\iff r\,(x^*, y^*) = 1 \iff x^* \text{ and } y^* \text{ have a positive linear relation, according to the property}$$

of the Pearson correlation.

2) Suppose $s\,(x, y) < 0$. Then $rm\,(x, y) = \dfrac{|s(x, y)|}{sm(x, y)} = \dfrac{-s(x, y)}{s(x^*, y')}$. So

$$rm\,(x, y) = r\,(x, y) \iff \frac{-s(x, y)}{s(x^*, y')} = \frac{s(x, y)}{s_x \cdot s_y} \iff s\,(x^*, y') = -s_x \cdot s_y \iff$$

$$\iff s\,(x^*, y') = -s_{x^*} \cdot s_{y'} \iff r\,(x^*, y') = -1 \iff x^* \text{ and } y' \text{ have a negative linear}$$

relation, according to the property of the Pearson correlation.

3) If $s\,(x, y) = 0$, then by the definitions $rm\,(x, y) = 0$ and $r\,(x, y) = 0$.   ■

Theorem 3 shows that the monotony and Pearson coefficients of $\begin{pmatrix} x \\ y \end{pmatrix}$ are equal when the ordered versions of x and y have a linear relation, and in the trivial case of no correlation.

5. The population monotony coefficient

In the previous sections we dealt with a sample measure of monotony. Here we briefly describe the population version of this measure.

In this section we fix a probability space. For random variables X and Y denote F_X the distribution function of X, $Cov\,(X, Y)$ the covariance and $\rho\,(X, Y)$ the

Pearson correlation coefficient of X and Y. Denote U a random variable with the uniform distribution on $[0, 1]$.

Definition 6. For the random variable X denote $X^* = F_X^{-1}(U)$ and $X' = F_X^{-1}(1 - U)$.

$\bullet$

The random variables X, X^* and X' have the same distribution function.

Definition 7. Suppose X and Y are random variables.

1) The **monotone covariance Covm** of X and Y is defined by:

$$Covm\,(X,\,Y) = \begin{cases} Cov(X^*,Y^*) \text{ if } Cov(X,Y) > 0, \\ 0 \text{ if } Cov(X,Y) = 0, \\ Cov(X^*,Y') \text{ if } Cov(X,Y) < 0. \end{cases}$$

$\bullet$

2) The **monotony coefficient ρm** of X and Y is defined by:

$$\rho m(X,Y) = \begin{cases} 0 \text{ if } Cov(X,Y) = 0, \\ \dfrac{|Cov(X,Y)|}{Covm(X,Y)} \text{ if } Cov(X,Y) \neq 0. \end{cases}$$

$\bullet$

Thus, the population and sample versions of the monotony coefficient are defined in similar ways.

Theorem 4. Relation to the population Pearson correlation. For the random variables X and Y the following holds.

1) If $Cov\,(X,\,Y) > 0$, then

$\rho m\,(X,\,Y) = \rho\,(X,\,Y) \quad \Leftrightarrow \quad$ there exist numbers a and $b > 0$, such that for all $x \in \mathbf{R}$,

$$F_X\,(x) = F_Y\,(a + bx).$$

2) If $Cov\,(X,\,Y) < 0$, then

$\rho m\,(X,\,Y) = \rho\,(X,\,Y) \quad \Leftrightarrow \quad$ there exist numbers a and $b < 0$, such that for all $x \in \mathbf{R}$,

$$F_X\,(x) = 1 - F_Y\,(a + bx).$$

3) If $Cov\,(X,\,Y) = 0$, then $\rho m\,(X,\,Y) = \rho\,(X,\,Y) = 0$. $\blacksquare$

In particular, Theorem 4 implies the equality of the monotony and Pearson coefficients when each of the variables X and Y has a normal distribution. This is consistent with one of the seven postulates stated in [8] for a measure of dependence of two random variables.

For the uniform distribution on the unit triangle these coefficients are different as the following example shows.

Example 4. Suppose the joint density function of the random variables X and Y is

given by: $\quad f(x, y) = \begin{cases} 2 \text{ if } 0 \leq x \leq y \leq 1, \\ 0 \text{ otherwise.} \end{cases}$

Then the Pearson correlation $\rho\,(X,\,Y) = 0.5$ and the population monotony coefficient $\rho m\,(X,\,Y) = \dfrac{2}{32 - 9\pi} \approx 0.5368.$ ∎

Example 5. Suppose X and ε are independent random variables, each with the standard normal distribution, and $Y = (X + \varepsilon)^3$.

Then $\rho\,(X,\,Y) = \dfrac{3}{\sqrt{30}} \approx 0.5477$ and $\rho m\,(X,\,Y) = \dfrac{1}{\sqrt{2}} \approx 0.7071.$ ∎

6. Discussion

The Pearson's correlation coefficient measures only linear dependence. In order to apply it, one needs to check the scatterplot for the data first. For example, if one is looking for relationships between 50 random variables, there are (50×49)/2 = 1225 scatterplots to check, which is quite time consuming. Besides conclusions from a scatterplot can be subjective: some people will see a linear relationship and others will not.

To apply the monotony coefficient rm, one does not need to look at the scatterplot. In the example with 50 random variables it is enough to compare 1225 numbers and select the ones with large absolute values (a computer can quickly do that); they will correspond to the pairs of random variables with high degrees of monotone dependence. The monotony coefficient applies to a larger class of dependencies than the class of linear dependencies.

The monotony coefficient is a tool that can be applied to samples in order to find dependencies between random variables; it is especially useful in finding couples of dependent variables in a big dataset of many variables. It helps stating hypotheses about data but not testing hypotheses, since the distribution of the monotony coefficient is unknown.

Comparison of the monotony coefficient rm with the Spearman's and Kendall's correlation coefficients shows that rm more accurately describes the degree of monotony in a relationship and it is more appropriate for continuous data, since it does not use ranks.

The population monotony coefficient $\rho m\,(X,\,Y)$ is also more appropriate for continuous random variables. It is defined in a more constructive way than the measures $\rho*$ and m_{xy} described in the introduction. For some bivariate distributions ρm can be calculated explicitly as in Examples 4 and 5. Such explicit calculation is not possible when the product $F_X^{-1}(u) \cdot F_Y^{-1}(u)$ cannot be integrated explicitly. In this case it still can be integrated by numerical methods and ρm can be evaluated approximately.

The authors are planning to do further research on the population monotony coefficient and present the results in next paper, which will potentially describe properties of the population monotony coefficient (similar to the ones in Theorem 2), more examples of the monotony coefficients for bivariate distributions and proofs of Theorem 4 and Examples 4, 5.

References

[1]   H. Gebelein, *Das statistiche problem der correlation als variations und eigenwertproblem und zusammenhang mit der ausgleichungsrechnung*, Zeitschrift für Angewandte Mathematik und Mechanik, Berlin: Akademie Verlag (1941).

[2]   G.Kimeldorf and A.R.Sampson, *Monotone dependence*, Annals Stat. 6, No. 4 (1978), pp. 895-903.

[3]   L.S.Mayer, *Estimating a correlation coefficient when one variable is not directly observed*, J. Am. Stat. Assoc. 68, No. 342 (1973), pp. 420-421.

[4]   M.Kendall and D.G.Gibbons, *Rank Correlation Methods,* 5th ed., Griffin, 1990.

[5]   F. Kachapova and I. Kachapov, *Monotone covariance of finite samples*, Mathematics and Informatics Quarterly 18, No. 3/4 (2008), pp. 81-91.

[6]   K. Schürger, *Wahrscheinlichkeitstheorie,* Oldenbourg, München − Wien, 1998.

[7]   K.D.Schmidt, *On the covariance of monotone functions of a random variable.* (2003). Available at
www.math.tu-dresden.de/sto/schmidt/dsvm/dsvm2003-4.pdf

[8]   A.Rényi, *On measures of dependence*, Acta Math. Acad. Sci. Hungar. 10, (1959), pp. 441-451.