Folia Oeconomica
Acta Universitatis Lodziensis
ISSN 0208-6018 e-ISSN 2353-7663

https://www.czasopisma.uni.lodz.pl/foe/
3(376) 2026
DOI: https://doi.org/10.18778/0208-6018.376.02

Martyna Kosińska https://orcid.org/0000-0002-5430-227X
University of Economics in Katowice, Department of Statistics, Econometrics and Mathematics, Katowice, Poland, martyna.kosinska@ue.katowice.pl

The Significance of Dependency for a Set of *k* Variables

Abstract: The issue of conducting multiple statistical tests on the same dataset is widely discussed in the statistical literature. As the number of hypotheses tested increases, so does the probability of committing a type I error. The aim of this article is to present a proposal for a permutation test that allows for testing the significance of the relationship between *k* variables as a whole. The application of this test yields a single hypothesis test, allowing the type I error to be controlled at the accepted significance level α. The article ends with an empirical analysis based on data from 2014 to 2023. For the years 2014–2017, the existence of statistically significant dependencies were confirmed for a set of three variables (gini coefficient, GDP *per capita* and unemployment rate).

Keywords: permutation tests, multiple comparisons, relationship, statistical significance

JEL: C12, C14, C15

Funding information: University of Economics in Katowice, Department of Statistics, Econometrics and Mathematics, Katowice, Poland.
Declaration regarding the use of GAI tools: Not used.
Conflicts of interests: None.
Ethical considerations: The Author assures of no violations of publication ethics and takes full responsibility for the content of the publication.
Received: 2026-01-21. Revised: 2026-04-29. Accepted: 2026-09-03

1. Introduction

In economical, social, medical studies and other disciplines, there is a need to perform multiple comparisons across groups simultaneously. The performance of many simultaneous statistical tests is connected with the multiple hypotheses testing problem. With each successive verified hypothesis, the probability of committing a type I error increases, which consists in rejecting the true hypothesis H_0. Various methods have been proposed in the literature to address this issue, including the Bonferroni and Holm methods. This article aims to verify the significance of the dependency of *k* variables, considering the dependency of a set of *k* variables, not variables individually.

Existing dependencies between individual variables may not be significant, but a set of *k* variables as a whole may create significant dependence. An empirical analysis using the proposed method was conducted for macroeconomic data taken from the Eurostat database (Gini coefficient, GDP *per capita*, and unemployment rate).

2. Significance of Pearson's Linear Correlation Coefficient

The verification of the significance of the dependency between two variables is based on the Pearson linear correlation coefficient test. The null hypothesis $H_0 : \rho = 0$ is verified against the alternative hypothesis $H_1 : \rho \neq 0$ ($\rho > 0$ or $\rho < 0$). For samples taken from a population with a two-dimensional normal distribution with a size of $n \leq 100$, assuming hypothesis H_0, the test statistic (Fisher, 1925) decribed by formula (1):

$$t = \frac{r}{\sqrt{1 - r^2}} \sqrt{n - 2} \tag{1}$$

has a Student's t-distribution with *n* – 2 degrees of freedom.

In Table 1, the minimum value of Pearson's linear correlation coefficient is presented, which leads to the recognition of dependence as a significant one for a given sample size, at a given significance level α.

Table 1. Critical values r_α for given *n* and α

n	α		
	0.01	**0.05**	**0.10**
5	0.9587	0.8783	0.8054
10	0.7646	0.6319	0.5494
15	0.6411	0.5139	0.4409
20	0.5614	0.4438	0.3783
27	0.4869	0.3809	0.3233
37	0.4182	0.3246	0.2746

n	*α*		
	0.01	**0.05**	**0.10**
47	0.3721	0.2875	0.2428

Source: own elaboration based on Krysicki et al., 1999.

3. Type I and Type II Error

The verification of statistical hypotheses is intrinsically linked to the possibility of making an incorrect decision. In the process of hypothesis verification, a distinction is made between two types of errors: type I error and type II error. The first one involves rejecting the null hypothesis when in fact it is true. The type II error is an error made when it is said that there are no grounds for rejecting the null hypothesis when it is false, and thus it is an error consisting in failing to detect a given effect. Danel (2016) states that it is more disadvantageous to reject the null hypothesis when it is true – that is, to commit a type I error.

With type I error, the concept of test size should also be taken into account. The test size is defined as a maximal (real) probability of rejecting the true null hypothesis (Holm, 1979). A single statistical test, when certain assumptions are met, ensures that the test size will not exceed the assumed significance level α. Differently, it is in the case of multiple hypothesis testing.

In practice, researchers can often encounter multiple hypothesis testing. It involves performing multiple tests on one dataset. Then the probability of making a type I error increases with each subsequent test (Danel, 2016). It is much more common for the true null hypothesis to be incorrectly rejected.

In studies, it is usually assumed that the maximum risk of rejecting the true null hypothesis can be equal to 0.05. The probability of committing one or more type I errors (rejecting at least one true null hypothesis) while performing multiple comparisons (*k* tests) can be described with a formula $1-(1-\alpha)^k$, where α is a significance level (Zar, 2010). In the case of performing 14 tests, the test size exceeds 0.5.

4. Selected Methods for Controlling Type I Error in Multiple Hypothesis Testing

Many methods have been developed to allow to control the type I error. Some of those are Bonferroni, sequential Bonferroni-Holm, Tukey, and Hochberg methods (Danel, 2016; Verma, Abdel-Salam, 2019).

In the Bonferroni method (Shaffer, 1995), the significance level α and the number of performed statistical tests *k* must be defined. Each *p*-value obtained during hypothesis testing has to be compared to $\frac{\alpha}{k}$ value (Bonferroni correction). If the *p*-value is less than or equal to

$\frac{\alpha}{k}$, the decision is made to reject the null hypothesis. Otherwise, there is no ground to reject it. The Bonferroni method is conservative, which means that using it makes it harder to obtain a statistically significant result.

The Bonferroni-Holm sequential method is a modification of the Bonferroni method and is a less conservative alternative. Testing *n* hypotheses is being considered. For each of the hypotheses $H_{01}, H_{02}, \ldots, H_{0n}$ *p*-value is assigned (Holm, 1979). *P*-values are sorted in an increasing order. The smallest *p*-value is compared to $\frac{\alpha}{k}$. If $p \leq \frac{\alpha}{k}$, H_0 is rejected in favour of H_1. Otherwise, there are no grounds to reject the null hypothesis and the algorithm stops. No significant effects (differences, dependencies, etc.) are found for the remaining hypotheses. If the null hypothesis is rejected in favour of the alternative hypothesis, the next smallest *p*-value is selected for comparison. Its value is considered with $\frac{\alpha}{k-i+1}$ value (where *i* is the iteration number). The algorithm is performed until the stop criteria are met or until the last hypothesis is rejected.

5. Simultaneous Verification of Dependencies for a Set of *k* Variables

In multiple hypothesis testing, a specified number of hypotheses are tested in a sequence. Stapor and Kończak (2025) have proposed a method that allows the type I error to be maintained at a predetermined level α. To achieve this goal, they used a sequence of permutation tests. In this proposition, based on a given solution, the goal is to state using only one test whether there are any dependencies among *k* variables. Each relationship is not considered separately, but rather the significance of the dependency within a set of variables. The aim is to determine whether certain relationships exist in a given set of variables.

The proposed method refers to the verification in the permutation Fisher-Pitman model (Kończak, 2020). Berry, Johnston, and Mielke (2014; 2018), and Berry et al. (2021) indicate that these methods are suitable for analysing data that are not necessarily random samples taken from a specific population. Further differences include no dependence on assumptions related to parametric tests, such as normality or homogeneity of variance, and the possibility of using it with non-random samples, which are so common in many studies. Moreover, Kończak (2016) describes permutation tests as those that depend only on the observed data and emphasises their usefulness in studies with small samples.

To verify the significance of the dependency for the set of *k* variables (X_1, X_2, ..., X_k), a test is proposed based on a correlation matrix *R*. In the case of independent variables, this matrix is the identity matrix *I*. The algorithm of the proposed test was described for the case in which dependencies are examined in a set of *k* variables as follows (Efron, Tibshirani, 1993; Good, 2005):

Formulating hypotheses and setting a significance level α.

$$H_0 : \rho = I \text{ (variables are independent)},$$

$$H_1 : \rho \neq I \text{ (there are dependencies between variables)}$$

where:
ρ – correlation coefficient matrix
I– unit matrix

$$\rho = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1k} \\ \rho_{21} & 1 & \ldots & \rho_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k1} & \rho_{k2} & \cdots & 1 \end{bmatrix} \mathbf{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Choice of the test statistic T (specified by formula 2).

$$T = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} |r_{ij}|, \tag{2}$$

where:
r_{ij} – linear correlation coefficient for variables X_i, X_j, $i, j = 1, 2, \ldots, k, i < j$.

Calculation of the test statistic value T_0 for the original dataset *N*-times (*N* = 1000) permutation of variables (assuming H0) and calculation of the test statistic value for each permutation $T_i (i = 1, 2, \ldots, N)$.

Calculation of the ASL value assessment described in formula (3):

$$\text{ASL} = P_{H_0}(T \geq T_0) \approx \frac{\text{card}\{i : T_i \geq T_0\}}{N}. \tag{3}$$

6. Making a decision

6.1. Multiple Dependency Testing – Empirical Analysis

In the analyses, one of the variables considered will be the Gini coefficient, a measure of income inequality. They are defined as unequal incomes of individuals and groups within one society. Income inequalities can be significant, which is undesirable. Highly unequal incomes can become a source of social unrest and a sense of injustice, leading to protests or rebellions against the authorities. At the other end of the spectrum, equal incomes for all – although an unrealistic situation – or low levels of income inequalities are also not expected. In such conditions, a sense of injustice may also arise – after all, every profession is different and has a different level of responsibility and risk – as well as stagnation, and reluctance to become more involved or to become an employee in certain professions. There is no optimal level of inequality. In principle, however, the aim of state policy is to reduce it.

A level of income inequality can be presented graphically by constructing a Lorenz curve (Figure 1). The OX axis shows the cumulative percentage of the population (households) and the OY axis presents the cumulative share of the *i*-th group in total income (Bellù, Liberati, 2005). The curve is created by connecting points with coordinates on the OX and OY axes. The closer the line is to a diagonal of the square formed (the egalitarian division line), the lower the level of income inequality.

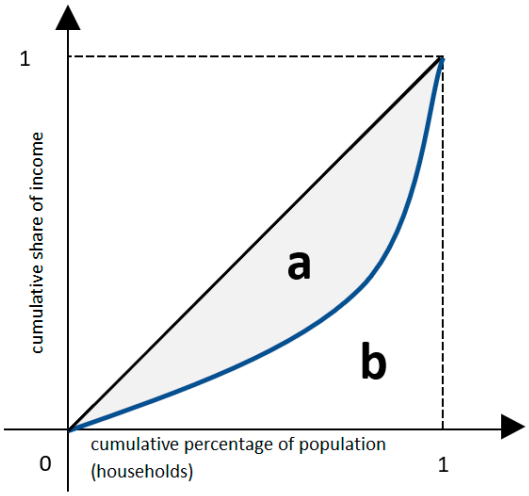

Figure 1. Lorenz curve
Source: own elaboration.

The Gini coefficient is used to measure the level of income inequality. It takes values from 0 to 1. The higher the index value (closer to 1), the greater the income inequality. The closer the index value to 0, the lower the inequalities. Zero describes a state in which incomes in groups are equal. The formula (4) presents how to calculate the Gini coefficient (Przekota, 2021):

$$G = \frac{a}{a+b} \in [0,1], \tag{4}$$

where:
a – the area under the egalitarian distribution curve (perfect equality) and the Lorenz curve,
b – the area under the Lorenz curve.

The Gini coefficient is a measure calculated on the basis of income or consumption expenditure (Haddad et al., 2024). In the following section, empirical analyses consider the Gini index based on equivalent disposable income. Equivalent disposable household income is defined as total income after tax and other charges, divided by comparable household size. It can be allocated to savings and expenditure on consumer goods and services. Household income is understood as the individual incomes of household members and the income allocated to the household. In order to determine the comparable size of the household, the following weights (Eurostat, n.d., *Glosary*...) are used:

1) 1 – the first adult in a household,
2) 0.5 – every other adult and a person aged 14 or over in a household,
3) 0.3 – each kid aged less than 14 in a household.

Assigning weights and summing them up allows to depicture differences in the size and composition of a household, enabling comparisons to be made, and in this case, comparable equivalent disposable income to be determined, and then on its basis, the Gini index to be calculated.

In the preliminary analysis of dependency of variables, there can be a question asked whether there are some significant correlations. To do this, the dataset from the Eurostat database (Eurostat, n.d., *Database*) was created, consisting of the five following variables.

1) *country*,
2) *year*,
3) *gini* – the Gini coefficient calculated on the basis of equivalent disposable income [in percentage],
4) *unemploy* – unemployment rate [in percentage],
5) *gdp* – GDP *per capita* [in USD].

The following figures show the development of the level of income inequalities (Figure 2), unemployment rate (Figure 3) and GDP *per capita* (Figure 4) in the European Union countries in 2023. The highest income inequalities occurred in Southern European countries (Bulgaria, Portugal, Spain, Italy and Greece), and Lithuania, Latvia and Estonia as well. On the other hand, some of the lowest inequalities were noted in Slovakia, the Czech Republic, Belgium, and Poland.

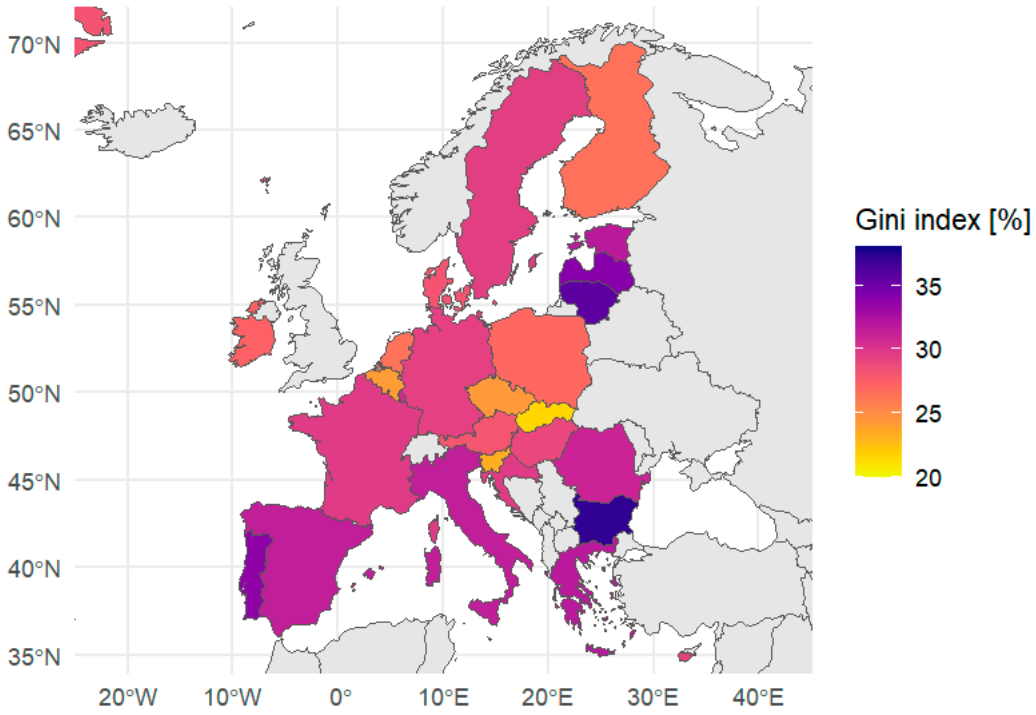

Figure 2. Income inequality in the European Union in 2023
Source: own elaboration based on data from Eurostat, n.d., *Database*.

In terms of unemployment, Greece and Spain had the highest rates, while Central European countries (Poland, the Czech Republic and Germany) as well as the Netherlands had the lowest. This seems to be determined by geography.

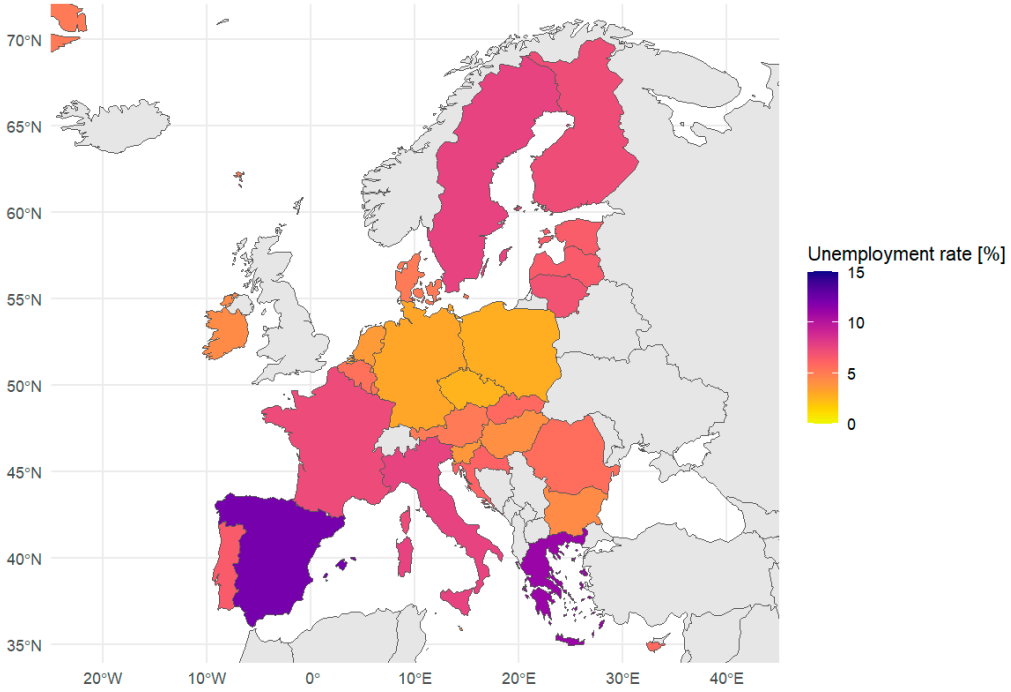

Figure 3. Unemployment rate in the European Union in 2023
Source: own elaboration based on data from Eurostat, n.d., *Database*.

The highest GDP *per capita* was recorded by Luxembourg and Ireland. At the same time, Figure 4 shows that Eastern European countries form a group of countries with the lowest GDP *per capita* values. Bulgaria recorded the lowest value.

The scatterplot matrix (Figure 5) was created from the presented variables. There was a negative linear correlation between the Gini coefficient and GDP *per capita*. The same was true for GDP *per capita* and the unemployment rate. A positive linear correlation was observed between the Gini coefficient and the unemployment rate. None of these correlations was strong. They were weak, and one of them was moderate at best. This is important for the proposed method due to the fact that none of the considered correlations would be recognised as statistically significant using a parametric test (even if the method's assumptions were met).

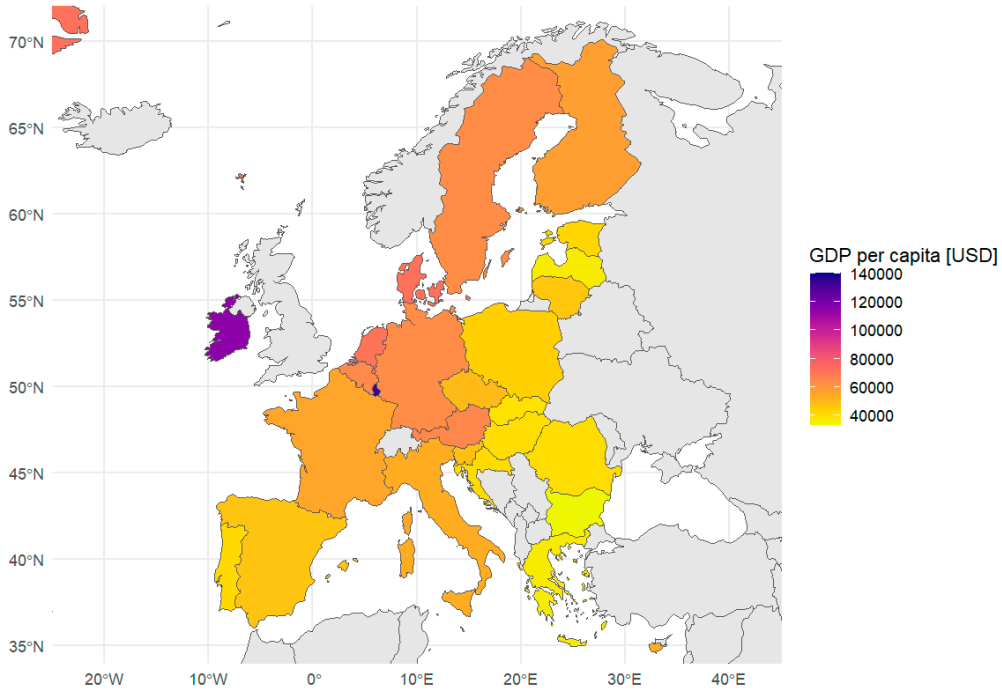

Figure 4. GDP *per capita* in the European Union in 2023
Source: own elaboration based on data from Eurostat, n.d., *Database*.

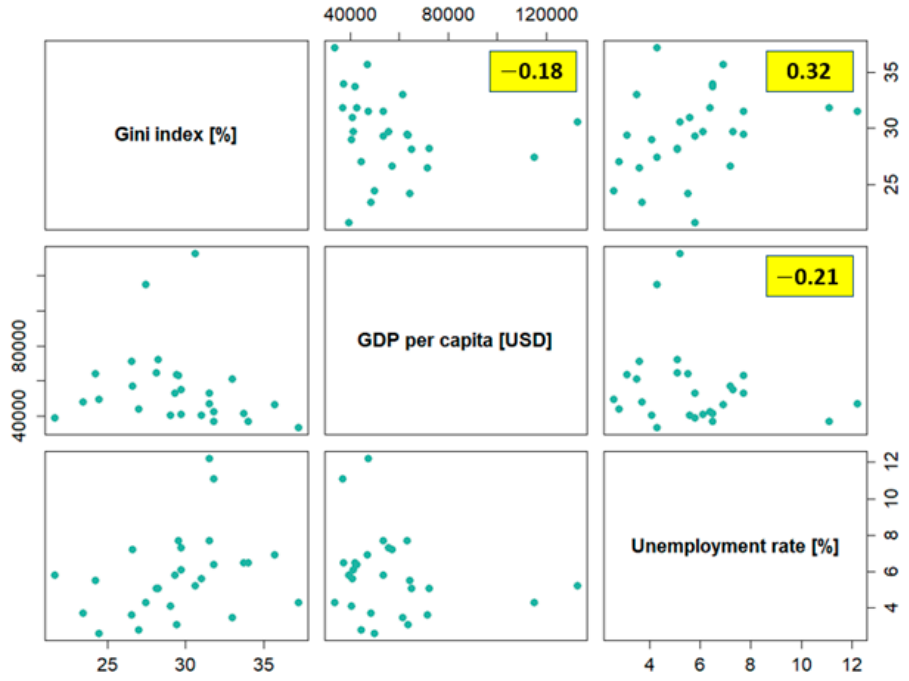

Figure 5. Gini coefficient, GDP *per capita* and unemployment rate in the European Union countries in 2017
Source: own elaboration.

Correlation coefficients between the considered variables in 2014–2023 (Table 2) confirm this view. The majority of these dependencies are weak or moderate at best. With 27 EU countries and 0.05 significance level, the Pearson linear correlation coefficient would be considered statistically significant if it were at least 0.3809. In Table 2, this would only be four positions.

Table 2. Pearson's linear correlation coefficient between the Gini coefficient, GDP *per capita* and the unemployment rate in 2014–2023

Year	Pearson's linear correlation coefficient		
	Gini – GDP *per capita*	GDP *per capita* – unemployment rate	Gini – unemployment rate
2014	– 0.385	– 0.377	0.487
2015	– 0.382	– 0.312	0.376
2016	– 0.318	– 0.274	0.423
2017	– 0.283	– 0.264	0.356
2018	– 0.188	– 0.209	0.291
2019	– 0.188	– 0.169	0.276
2020	– 0.167	– 0.186	0.312
2021	– 0.280	– 0.197	0.275
2022	– 0.243	– 0.240	0.302
2023	– 0.185	– 0.212	0.321

Source: own elaboration.

In empirical analyses, the proposed permutation test was used to analyse dependencies among the variables gini, gdp, and unemployment for the years 2014–2023 (27 European Union countries). Estimated ASL values were presented in Table 3. Correlations for the years 2014–2017 were statistically significant, but not for the other years.

Table 3. Estimated ASL

Year	Estimated ASL
2014	0.000
2015	0.005
2016	0.015
2017	0.029
2018	0.139
2019	0.205
2020	0.149
2021	0.102
2022	0.077
2023	0.117

Source: own elaboration.

Analysing the year 2017 is especially interesting. Correlations between the following variables: Gini index (var. *gini*), unemployment rate (var. *unemploy*) and GDP *per capita* (var. *gdp*) were calculated. Results were presented in a correlation matrix (Table 4). None of the correlations would be considered significant using a t-test of significance (with a sample size *n* = 27, the critical value $r_\alpha = 0.380$).

Table 4. Correlation matrix for three variables

Correlation	gini	unemploy	gdp
gini	1	0.356	0.283
unemploy	0.356	1	0.264
gdp	0.283	0.264	1

Source: own elaboration.

In order to verify the hypothesis $H_0 : \rho = I$ against $H_1 : \rho \neq I$, the proposed method was used. The empirical distribution of T_i statistic was obtained (Figure 6). The value of T_0 statistic is marked in red.

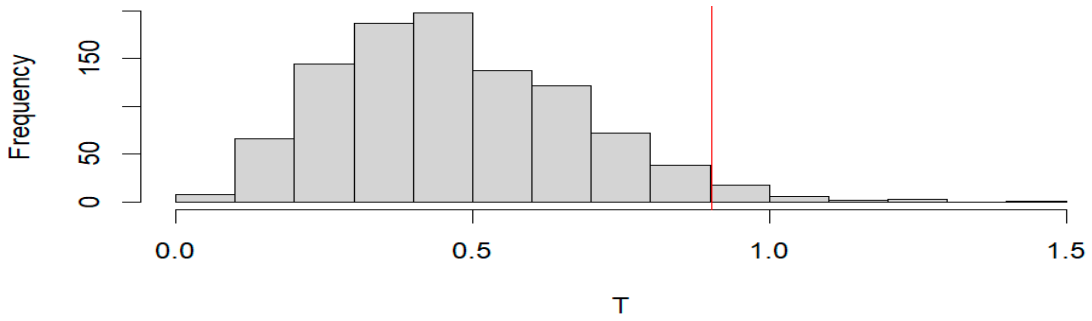

Figure 6. Empirical distribution of *T* statistic
Source: own elaboration.

Estimated ASL was equal to 0.029. It was less than the assumed significance level α. The null hypothesis is rejected in favour of the alternative hypothesis. There are certain dependencies between the variables.

7. Conclusions

In the article, a method was proposed for verifying the significance of dependencies among a set of *k* variables. The proposed method uses a permutation test and provides type I error control at the a level. A Monte Carlo study was used to estimate the ASL, and the calculations were based on proprietary procedures implemented in R.

The proposed method was used to verify the statistical significance of dependency for a set of three variables: gini coefficient, GDP *per capita* and unemployment rate. Analyses were conducted for 27 European Union countries for the years 2014–2023. The existence of statistically significant correlations was confirmed for the years 2014–2017. In particular, in 2017, none of the individual correlations were significant (even if the assumptions of multivariate normality were met), but the set of three variables showed statistically significant dependencies, not random ones.

The proposed method allows to verify existing dependencies for a set of *k* variables. It is a different way of looking at the issue of multiple hypothesis testing and constitutes an extension of statistical methodology. This method can be particularly useful for initial analyses and research focused on the occurrence of dependencies within a larger set of variables.

References

Bellù L.G., Liberati P. (2005), *Charting Income Inequality: The Lorenz Curve*, https://openknowledge.fao.org/server/api/core/bitstreams/7c681f66-90d8-4c4e-b006-8e6931bfb034/content [accessed: 23.09.2025].

Berry K.J., Johnston J.E., Mielke Jr. P.W. (2014), *A chronicle of permutation statistical methods: 1920–2000, and beyond*, Springer, Cham.

Berry K.J., Johnston J.E., Mielke Jr. P.W. (2018), *The Measurment of Association. A Permutation Approach*, Springer Nature Switzerland, Cham.

Berry K.J., Kvamme K.L., Johnston J.E., Mielke Jr. P.W. (2021), *Permutation Statistical Methods with R*, Springer International Publishing, Cham.

Danel D. (2016), *Poziom istotności i granica rozsądku – problem porównań wielokrotnych w badaniach naukowych na przykładach z zakresu biologii behawioralnej człowieka*, https://media.statsoft.pl/pdf/czytelnia/poziom_istotnosci_i_granica_rozsadku.pdf [accessed: 19.06.2026].

Efron B., Tibshirani R. (1993), *An Introduction to the Bootstrap*, Chapman & Hall/CRC, Boca Raton.

Eurostat (n.d.), *Database*, https://ec.europa.eu/eurostat/data/database [accessed: 19.06.2026].

Eurostat (n.d.), *Glossary: Equivalised disposable income*, https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary: Equivalised_disposable_income [accessed: 18.06.2026].

Fisher R.A. (1925), *Statistical Methods for Research Workers*, Oliver and Boyd, Edinburgh.

Good P. (2005), *Permutation, Parametric and Bootstrap Tests of Hypotheses*, Science Business Media, Inc., New York.

Haddad C.N., Mahler D.G., Diaz-Bonilla C., Hill R., Lakner C., Ibarra G.L. (2024), *The World Bank's New Inequality Indicator: The Number of Countries with High Inequality*, "Policy Research Working Paper", 10796, https://doi.org/10.1596/41687

Holm S. (1979), *A Simple Sequentially Rejective Multiple Test Procedure*, "Scandinavian Journal of Statistics", vol. 6(2), 2, pp. 65–70.

Kończak G. (2016), *Testy permutacyjne: Teoria i zastosowania*, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice.

Kończak G. (2020), *Nieklasyczne metody statystyczne w badaniach ekonomicznych*, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice.

Krysicki W., Bartos J., Dyczka W., Królikowska K., Wasilewski M. (1999), *Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach. Część II. Statystyka matematyczna*, Wydawnictwo Naukowe PWN, Warszawa.

Przekota G. (2021), *Wybrane koncepcje pomiaru nierówności dochodowych*, „Nierówności Społeczne a Wzrost Gospodarczy", no. 66(2), pp. 16–32, http://doi.org/10.15584/nsawg.2021.2.2
Shaffer J.P. (1995), *Multiple Hypotesis Testing*, "Annual Review Psychology", vol. 46, pp. 569–570.
Stapor K., Kończak G. (2025), *An Alternative Powerful Approach to Multiple Hypotheses Testing: Application in a Correlation Study*, "International Journal of Applied Mathematics and Computational Science", vol. 35(4), pp. 677–686, https://doi.org/10.61822/amcs-2025-0048
Verma J.P., Abdel-Salam A.-S.G. (2019), *Testing statistical assumptions in research*, John Wiley & Sons, Hoboken.
Zar J.H. (2010), *Biostatistical analysis*, Prentice-Hall/Pearson, Upper Saddle River.

Istotność zależności dla układu *k* zmiennych

Streszczenie: Zagadnienie przeprowadzania wielu testów statystycznych na tym samym zbiorze danych jest szeroko opisywane w literaturze statystycznej. Wraz ze wzrostem liczby weryfikowanych hipotez rośnie prawdopodobieństwo popełnienia błędu I rodzaju. Celem artykułu jest przedstawienie propozycji testu permutacyjnego, który pozwala na sprawdzenie istotności zależności układu *k* zmiennych jako całości. Zastosowanie tego testu prowadzi do weryfikacji tylko jednej hipotezy, co pozwala utrzymać kontrolę błędu I rodzaju na przyjętym poziomie istotności a. Artykuł kończy empiryczna analiza oparta na danych z lat 2014–2023. W latach 2014–2017 potwierdzono istnienie istotnych statystycznie zależności dla układu trzech zmiennych (współczynnik Giniego, PKB na mieszkańca i stopa bezrobocia).

Słowa kluczowe: testy permutacyjne, porównania wielokrotne, zależność, istotność statystyczna