Wiadomości Statystyczne. The Polish Statistician, 2025, *70*(5), 1–15 https://doi.org/10.59139/ws.2025.05.1
Studia metodologiczne / Methodological studies
Zgłoszony/submitted: 16.01.2024, poprawiony/revised: 20.11.2024, zaakceptowany/accepted: 3.03.2025

# Use of permutation methods in testing the independence of variables in multidimensional contingency tables[1]

Grzegorz Kończak,[a] Martyna Kosińska[b]

**Abstract.** Multiple categories of qualitative variables are often identified in economic and social research. In these cases, contingency tables (two- or multidimensional) are frequently used to report the results of a study. Statistical inference based on data provided in contingency tables involves in most instances the use of the chi-square test of independence or the chi-square test of homogeneity. The form of the statistics in both tests is the same, but the method by which the data are obtained remains different. A sample of size $n$ is drawn from a single population with two classification variables in the first case, and in the second samples from an $r$ population of $n_1, n_2, \dots, n_r$ sizes are drawn independently, considering a single classifier variable only. It is much less common to make comparisons of structures for more than two qualitative variables simultaneously. The aim of the research described in this article is to present the usefulness of permutation methods in testing the presence of significant differences in the structures of the levels of participation in cultural events based on data contained in multidimensional contingency tables. The application of the permutation test was illustrated through the results of the authors' own research examining the cultural participation of active users of Internet portals immediately before and during the COVID-19 pandemic. Participation in cinema screenings by gender, age and work status was analysed. The research results confirmed that the pandemic had a significant impact on participation in cinema screenings.
**Keywords:** statistical inference, multidimensional contingency tables, permutation tests
**JEL:** C12, C15, C18

# Zastosowanie metod permutacyjnych w testowaniu niezależności zmiennych w wielowymiarowych tablicach kontyngencji

**Streszczenie.** W badaniach gospodarczych i społecznych niejednokrotnie wyróżnia się wiele kategorii zmiennych jakościowych. W takich przypadkach do raportowania wyników badania często stosuje się tablice kontyngencji (dwu- lub wielowymiarowe). Wnioskowanie statystyczne na

[1] Artykuł został opracowany na podstawie referatu wygłoszonego na 41st International Conference MSA'2023, która odbyła się w dniach 6–8 listopada 2023 r. w Łodzi. / The article is based on a paper delivered at the 41st International Conference MSA'2023 held on 6–8 November 2023 in Lodz, Poland.

[a] Uniwersytet Ekonomiczny w Katowicach, Wydział Zarządzania, Polska / University of Economics in Katowice, Faculty of Management, Poland. ORCID: https://orcid.org/0000-0002-4696-8215. Autor korespondencyjny / Corresponding author, e-mail: grzegorz.konczak@ue.katowice.pl.

[b] Uniwersytet Ekonomiczny w Katowicach, Wydział Zarządzania, Polska / University of Economics in Katowice, Faculty of Management, Poland. ORCID: https://orcid.org/0000-0002-5430-227X. E-mail: martyna.kosinska@ue.katowice.pl.

podstawie danych zawartych w takich tabelach zazwyczaj przeprowadzane jest z wykorzystaniem testu niezależności chi-kwadrat lub testu jednorodności chi-kwadrat. W obu testach postać statystyki jest taka sama, ale mechanizm powstawania danych – odmienny. W pierwszym przypadku próba o rozmiarze $n$ jest losowana z pojedynczej populacji z dwiema zmiennymi klasyfikacyjnymi. W drugim – próbki z $r$ populacji o rozmiarach $n_1, n_2, \dots, n_r$ są losowane niezależnie, uwzględniając tylko jedną zmienną klasyfikującą. Porównania struktur dla trzech i więcej zmiennych jakościowych są znacznie rzadsze niż dla dwóch zmiennych. Celem artykułu jest ukazanie przydatności metod permutacyjnych w testowaniu istotności różnic w strukturach na podstawie danych zawartych w wielowymiarowych tablicach kontyngencji. Zastosowanie testu permutacyjnego pokazano na przykładzie własnego badania dotyczącego uczestnictwa w kulturze aktywnych użytkowników portali internetowych bezpośrednio przed pandemią COVID-19 i w jej trakcie. Analizowano uczestnictwo w seansach kinowych w zależności od płci, wieku i statusu zawodowego. Wyniki badania potwierdziły istotny wpływ pandemii na uczestnictwo w seansach kinowych.

**Słowa kluczowe:** wnioskowanie statystyczne, wielowymiarowe tablice kontyngencji, testy permutacyjne

## 1. Introduction

In statistical research, the inference of unknown population parameters is often carried out on the basis of random samples. In the case of characteristics measured on strong scales (interval and ratio), samples of relatively small sizes are sufficient for an effective statistical inference. The most common tests are Student's $t$-test for comparing two populations and the ANOVA test for comparing $r \ (r > 2)$ populations if the assumption of normality of distribution is met. If not, then tests based on the Mann-Whitney $U$ rank analysis and the Kruskal-Wallis test are most often used. When inferring from characteristics measured on weak scales (ordinal or nominal), independence or chi-square homogeneity tests are most frequently used. These tests, however, require relatively larger samples than tests used for variables on strong scales. Inference from two- and multidimensional[2] contingency tables data requires at least 5 counts in all cells of the array (Sheskin, 2004; Zeliaś et al., 2002). If this condition is not met, then, for example, for $2 \times 2$ arrays methods that incorporate the Yates and Dandekar corrections (Rao, 1973; Yates, 1934) or Fisher's exact test (Fisher, 1935) can be used. Due to computational complexity, the literature considers the exact test for arrays of $2 \times 2$. Currently, the available software allows the calculation of the corresponding probabilities even for arrays of larger dimensions. The problem increases when the number of cells in the contingency table is large, i.e. when the expected frequencies are small. Along with the availability of computational techniques, it is also possible to use simulation approximations of the exact test. The aim of the research described in this article is to present the usefulness of permutation methods in testing the presence of significant differences in the structures of the levels

[2] In this article, the term multidimensional (multivariate, multiway or multi-way) contingency table is used to describe a table that is designed to analyse the relationships between three or more categorical variables.

of participation in cultural events based on data contained in multidimensional contingency tables. The study highlights the particular utility of permutation methods for such analyses, primarily due to their non-parametric nature, which circumvents the restrictive distributional assumptions often unsuitable for complex categorical data.

## 2. Inference for two-dimensional contingency tables: chi-square and exact tests

In economic and social scientific research, results are often presented in contingency tables. There are two chi-square tests used for data in contingency tables: the chi-square test of independence and the chi-square test for homogeneity (Domański, 1979; Sheskin, 2004). The chi-square test of independence is applied when a single sample is categorised into two variables. It is assumed that the sample is randomly selected from the population it represents. This test evaluates the general hypothesis that the two variables are independent from each other. The chi-square test for homogeneity applies when $r$ independent samples (where $r \geq 2$) are categorised on one dimension only, i.e. by one variable which for a given unit can take one of the $c$ categories (where $c \geq 2$). Data for $r$ independent samples are recorded for the number of observations in each sample that fall into all of the $c$ categories. Each sample is assumed to be randomly drawn from the population. The chi-square test for homogeneity evaluates whether the $r$ samples are homogeneous in terms of the proportion of observations in each of the $c$ categories. More specifically, if the data are homogeneous, the proportion of observations in the $j$-th category will be equal in all $r$ populations.

Both tests, the one that checks for independence and the one that checks for homogeneity, are based on certain assumptions (Sheskin, 2004):

a) categorical/nominal data (i.e. frequencies) for $r \times c$ mutually exclusive categories are employed in the analysis;
b) the data that are evaluated represent a random sample comprised of $n$ independent observations. This assumption reflects the fact that each subject or object can be represented in the data only once;
c) the expected frequency of each cell in the contingency table is 5 or greater. When the expected frequency of one or more cells is less than 5, the probabilities in the chi-square distribution may not provide an accurate estimate of the underlying sampling distribution. As is the case for the chi-square goodness-of-fit tests, sources are not in agreement with respect to the minimum acceptable value for an expected frequency.

For both chi-square tests, the form of the test statistic is the same and has the following form (Domański, 1979, p. 160; Sheskin, 2004, p. 520; Zeliaś et al., 2002, p. 110):

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\left(n_{ij} - \hat{n}_{ij}\right)^2}{\hat{n}_{ij}}, \tag{1}$$

where:
$n_{ij}$ – observed frequencies,
$\hat{n}_{ij}$ – expected frequencies under the null hypothesis (H0), conditionally on the margins of the contingency table.

In the following part of the article, we will consider a hypothesis examining independence between three or more variables, which can be expressed as:

- H0: all variables are mutually independent;
- H1: at least one association exists between the variables.

The statistic given by (1) under H0 has the asymptotic chi-square distribution with $df = (r-1)(c-1)$ degrees of freedom. Some authors believe that the quantiles of the chi-square distribution for the test of independence can only be used if the number of observations in each cell of the contingency table is at least equal to 5, 8 or 10 (Domański, 1979; Fisz, 1967; Greń, 1974; Siegel, 1956). When the expected frequencies do not satisfy this rule, then the chi-square approximation is not appropriate, but we can instead use an exact test or develop a permutation test.

For small samples, when hypothesis H0 is true, the rejection rate of H0 using the Pearson chi-square test can be too high (Little, 1989). For 2 × 2 tables (where $r = 2$ and $c = 2$), there are some alternatives for the chi-square test such as the chi-square test with the Yates or Dandekar correction (Rao, 1973).

Little (1989) presents a table that shows the percentage of rejection rates under H0 for nominal significance level $\alpha = 0.05$ (see Table 1). Analysing the results from this table, we can conclude that Fisher's exact test is much too conservative in small samples.

**Table 1.** Percentage of rejection rates under H0 of Pearson and Fisher tests for nominal $\alpha = 0.05$

| $n_1 = n_2$ | $p_0 = p_1 = 0.5$ | | $p_0 = 0.2,\ p_1 = 0.8$ | |
|---|---|---|---|---|
| | Pearson | Fisher | Pearson | Fisher |
| 5 | 5.5 | 1.1 | 2.2 | 2.0 |
| 10 | 5.8 | 2.1 | 4.6 | 1.5 |
| 20 | 4.2 | 2.1 | 5.1 | 2.3 |
| 50 | 4.5 | 2.9 | 5.0 | 2.9 |
| 100 | 5.2 | 3.8 | 4.9 | 3.5 |
| 200 | 4.9 | 4.0 | . | . |
| Infinity | 5.0 | 5.0 | 5.0 | 5.0 |

Note. $p_0$, $p_1$ – theoretical proportions in two groups for a 2 × 2 contingency table.
Source: Little (1989).

It is clear from Table 1 that, assuming H0 is true, both the Pearson chi-square test and the Fisher exact test have rejection rates close to $\alpha$ only for large samples. The rates for small samples are significantly higher than the $\alpha$ nominal value, especially when performing the Pearson chi-square test. Sometimes, instead of the Fisher exact test, it is better to use the Boschloo exact test (Berger & Boos, 1994), which is based on the $p$-values for Fisher's exact test. Agresti and Coull (1998) discuss a similar problem that appears during the construction of confidence intervals on proportions. They recommend using approximations of distributions instead of exact distributions. Following this concept, we use the test in inference for data in multidimensional contingency tables. This is based on the Monte Carlo approximation of the exact test for multidimensional tables.

## 3. Extension of the chi-square test for multidimensional contingency tables

The chi-square test for $r \times c$ tables can be generalised for use with multidimensional contingency tables, containing information on three or more variables. They could be represented as a cross-tabulation of the frequencies of different combinations of categories across the variables. The variables could be nominal, ordinal, interval or ratio. Interval or ratio variables have to be categorised for analysis in (multidimensional) contingency tables.

In a three-dimensional table, one of the variables is designated as the row variable, the second as the column variable, and the third as the layer variable (Sheskin, 2004, p. 571). Christensen (1990, pp. 63–64) notes that among the possible ways of conceptualising the relationship between the variables in a three-dimensional contingency table are the following:

a) rows, columns and layers are all independent from one another;
b) rows are independent from the columns and layers (with columns and layers not necessarily being independent from one another);
c) columns are independent from the rows and layers (with rows and layers not necessarily independent from one another);
d) layers are independent from the rows and columns (with rows and columns not necessarily independent from one another);
e) given any specific level of a row, columns and layers are independent from one another;
f) given any specific level of a column, rows and layers are independent from one another;
g) given any specific level of a layer, rows and columns are independent from one another.

In the case of a), the null and alternative hypothesis for the three-dimensional contingency table could be written in the following form:

- H0: in the underlying population, the three variables are all independent from one another;
- H1: in the underlying population, the three variables are not all independent from one another.

The test statistics in the three-dimensional case takes the following form:

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\sum_{k=1}^{l}\frac{\left(n_{ijk} - \hat{n}_{ijk}\right)^2}{\hat{n}_{ijk}}, \tag{2}$$

where:
$r$ – number of rows,
$c$ – number of columns,
$l$ – number of layers in the contingency table,
$n_{ijk}$ – observed frequencies,
$\hat{n}_{ijk}$ – expected frequencies.

Statistic (2) under H0 has an asymptotic chi-square distribution (Sheskin, 2004) with $df = rcl - r - c - l + 2$ degrees of freedom.

Loglinear models (Agresti & Gottard, 2007; Simonoff, 2003), exact test approximation (Cheon, 2012) and graphical methods (Friendly, 1994; Stanimir, 2011) are used to analyse dependencies between variables for data in multidimensional contingency tables. Perrone et al. (2024) proposed transforming a multidimensional contingency table into a table with uniform margins and the same dependence structure. Sheskin (2004) points out that as the number of variables (as well as the number of levels of categories per variable) increases, the analysis of a multidimensional table becomes increasingly complex. In addition, the more variables included in a study, the more difficult it becomes to obtain a clear interpretation of the results of any analyses that are conducted. As the number of cells in a multidimensional contingency table increases, more and more samples are needed to fulfil the requirement for a minimum expected count of 5.

**Table 2.** Number of cells and the necessary sample size for a chi-square test for some three-dimensional contingency tables

| Variable | Scenario 1 | Scenario 2 | Scenario 3 |
|---|---|---|---|
| $X$ | 2 | 2 | 4 |
| $Y$ | 2 | 3 | 5 |
| $Z$ | 2 | 4 | 5 |
| Number of cells | 8 | 24 | 100 |
| Necessary sample size | 40 | 120 | 500 |

Source: authors' work.

**Table 3.** Number of cells and the necessary sample size for a chi-square test for some four-dimensional contingency tables

| Variable | Scenario 1 | Scenario 2 | Scenario 3 |
|---|---|---|---|
| $X$ | 2 | 2 | 4 |
| $Y$ | 2 | 3 | 5 |
| $Z$ | 2 | 4 | 5 |
| $W$ | 2 | 5 | 6 |
| Number of cells | 16 | 120 | 600 |
| Necessary sample size | 80 | 600 | 3,000 |

Source: authors' work.

Tables 2 and 3 show the number of cells in some multidimensional contingency tables and the necessary sample size required to ensure minimum expected counts of 5 in all cells of each table. Table 2 gives such examples (scenarios) for selected three-dimensional contingency tables and Table 3 for selected four-dimensional contingency tables. In the case of the three-dimensional contingency table (Table 2), where each variable can take 2 variants for sample sizes of less than 40, it is not possible to ensure all expectation counts of 5. A much larger sample size will be often necessary. Similarly, in the case of the four-dimensional contingency table (Table 3), when the variables can take 4, 5, 5, and 6 variants for sample sizes below 3,000, it is not possible to provide all the expected counts at level 5. A much larger sample size will be often necessary, in particular in the case of a substantial imbalance of observations across cells in the multidimensional contingency table. The permutation approach utilises exact and Monte Carlo resampling permutation statistical procedures to generate probability values for a variety of measures of association (Berry et al., 2018).

## 4. Permutation tests for multidimensional contingency tables

The use of permutation methods for precise inference dates back to the work of Fisher published in 1935. Since then, the practicality of such methods has increased steadily along with computing power (Ernst, 2004). Permutation tests are non-parametric tests that require very few assumptions. The general purpose of a permutation test is to estimate the test statistic distribution under H0. On this basis, we can determine how rare the observed frequencies are relative to the estimated distribution and decide which hypothesis to test (Berry et al., 2016; Good, 2005).

The previous section, following Christensen (1990), provided various examples of possible hypotheses for data in the form of a three-dimensional contingency table. Figures 1 and 2 show two selected methods for permuting data. Other methods for permuting multivariate data can be found in O'Gorman (2012). Figure 1 schematically shows the permutation of observations in terms of one variable (variable $X$).

**Figure 1.** Permutation of observations of one variable ($X$)

Sample data

| $X$ | $Y$ | $Z$ |
|---|---|---|
| $x_1$ | $y_1$ | $z_1$ |
| $x_2$ | $y_2$ | $z_2$ |
| $x_3$ | $y_3$ | $z_3$ |
| ... | ... | ... |
| $x_n$ | $y_n$ | $z_n$ |

Permuted data

| $X$ | $Y$ | $Z$ |
|---|---|---|
| $x_2$ | $y_1$ | $z_1$ |
| $x_n$ | $y_2$ | $z_2$ |
| $x_7$ | $y_3$ | $z_3$ |
| ... | ... | ... |
| $x_5$ | $y_n$ | $z_n$ |

| $X$ | $Y$ | $Z$ |
|---|---|---|
| $x_4$ | $y_1$ | $z_1$ |
| $x_3$ | $y_2$ | $z_2$ |
| $x_n$ | $y_3$ | $z_3$ |
| ... | ... | ... |
| $x_2$ | $y_n$ | $z_n$ |

Source: authors' work.

Figure 2 schematically shows the permutation within observations for variables $X$ and $Y$. Such permutation is possible only if variables $X$ and $Y$ are paired observations, e.g. it is a measurement of the same phenomenon in different periods.

**Figure 2.** Permutation of observations between variables $X$ and $Y$ for a fixed observation

Sample data

| $X$ | $Y$ | $Z$ |
|---|---|---|
| $x_1$ | $y_1$ | $z_1$ |
| $x_2$ | $y_2$ | $z_2$ |
| $x_3$ | $y_3$ | $z_3$ |
| ... | ... | ... |
| $x_n$ | $y_n$ | $z_n$ |

Permuted data

| $X$ | $Y$ | $Z$ |
|---|---|---|
| $x_1$ | $y_1$ | $z_1$ |
| $y_2$ | $x_2$ | $z_2$ |
| $x_3$ | $y_3$ | $z_3$ |
| ... | ... | ... |
| $y_n$ | $x_n$ | $z_n$ |

| $X$ | $Y$ | $Z$ |
|---|---|---|
| $y_1$ | $x_1$ | $z_1$ |
| $y_2$ | $x_2$ | $z_2$ |
| $x_3$ | $y_3$ | $z_3$ |
| ... | ... | ... |
| $y_n$ | $x_n$ | $z_n$ |

Source: authors' work.

Depending on the form of H0, an appropriate form of permutation should be chosen so that the empirical distribution of the test statistic approximates the theoretical distribution of this statistic under H0.

The permutation test procedure involves the following steps (Good, 2005; Kończak, 2012):

1. setting up the H0 and the alternative hypothesis;
2. assuming level of significance $\alpha$;
3. choosing the form of test statistic $T$;
4. calculation of the value of the test statistic ($T_0$) for the sample data;
5. permutation of the variable or variables $N$ times and for each permutation, the calculation of the $T_i$ statistic ($i = 1, 2, \ldots, N$) value;
6. making a decision to reject or not to reject H0 based on the value of the $T_0$ statistic and the empirical distribution of the $T$ statistic obtained by simulation.

The decision is made on the basis of the achieved significance level (ASL) value given by Efron and Tibshirani (1993):

$$ASL = P_{\mathrm{H0}}(T \geq T_0) \approx \frac{card\{i: T_i \geq T_0\}}{N}. \quad (3)$$

The smaller the value of ASL, the stronger the evidence against H0.

The chi-square statistic defined in equation (1) can serve as test statistic $T$, providing a quantitative measure of the degree of inconsistency occurring within the structure of the contingency tables. A test constructed this way as a statistical estimation of the exact test, with a sufficient number of simulation replications assures a test size at a significance level no greater than $\alpha$ (Berry et al., 2014; Kończak, 2016).

## 5. Cultural participation during the COVID-19 pandemic – inference from multidimensional contingency tables based on own research results

The economic and social impact of the COVID-19 pandemic is undeniable. It was characterised by diversity, as indicated in the Santander (2023) report. In some cases, it was positive, as e.g. in the case of courier services or the production of hygiene products. Other sectors such as those selling food were not significantly affected. On the other hand, economic sectors whose proper functioning relied on large numbers of people, e.g. the catering, education and cultural sectors, were severely challenged by the COVID-19 pandemic. Raevskikh et al. (2022) distinguish several aspects affected by the phenomenon, involving e.g. significant revenue losses reported by businesses, jeopardised jobs and livelihoods, increased consumption of cultural heritage at home, reduced access to cultural events, or changes in the way culture was created and distributed. Faced with changing operating conditions, cultural entities and cultural providers had to modify their approach. Technology stepped in, e.g. to broadcast concerts and performances online, to make archived theatre performances available, to digitise and make accessible library and museum collections, or even to provide virtual museum tours.

The survey on participation in cultural events before and during the COVID-19 pandemic was conducted in the first half of 2022. The survey questionnaire was made available on a major social network. Responses were given anonymously. The survey itself dealt with the respondents' behaviours and feelings towards the changing situation in the sphere of culture during the COVID-19 pandemic. Conducting the survey and analysing the results was aimed at finding out the respondents' opinions on the changes in their behaviours in the context of cultural use in the COVID-19 pandemic. Additionally, the aim was to confirm the impact of the pandemic on the cultural sector, also on the forms and frequency of cultural participation of active users of the social network.

The majority of the 203 respondents were women: there were 109 women and 94 men. In terms of age, most respondents were between 20 and 30 years old. The educational level of the respondents varied, with those holding an engineer's or bachelor's degree (58.6%) forming the largest group, or with a high school education (23.6%). As far as the structure of the respondents' professional status is concerned, it should be noted that almost half of them (41.9%) combined studies and work. The primary dataset consisted of 29 variables (Kończak & Kosińska, 2023; Kosińska, 2023). Six of them were used for further analysis:

- attendance at cinema screenings before the COVID-19 pandemic (Before.CV.19);
- attendance at cinema screenings during the COVID-19 pandemic (During.CV.19);
- gender (Gender);
- age (Age);
- status of the respondents (Status);
- degree of agreement with the sentence: TV and online movies have replaced my ability to go to the cinema (Degree.of.compliance).

Their names and their possible categories are presented in Table 4.

**Table 4.** Set of the used variables and their variants

| Variable | Variants of variables |
|---|---|
| Before.CV.19 | rarely, sometimes, often |
| During.CV.19 | rarely, sometimes, often |
| Gender | man, woman |
| Age | up to 24, over 24 |
| Status | study, study&work, work, other |
| Degree.of.compliance | agree, neutral, disagree |

Source: authors' work.

In this paper, we use a permutation test to examine the independence between the three variables for the following four sets:

- Before.CV.19, During.CV.19, Gender;
- Before.CV.19, During.CV.19, Age;
- Before.CV.19, During.CV.19, Status;
- Age, Status, Degree.of.compliance.

Figure 3 shows a mosaic graph (Friendly, 1994) that juxtaposes three variables: Before.CV.19, During.CV.19, and Gender. Each rectangle represents the count in the cell at the intersection of the three variable categories. The scale on the right indicates the levels of standardised residuals $d_{ijk}$. In the three-dimensional case, the residuals are as follows (Friendly, 1994, p. 190; Sheskin, 2004, p. 546):

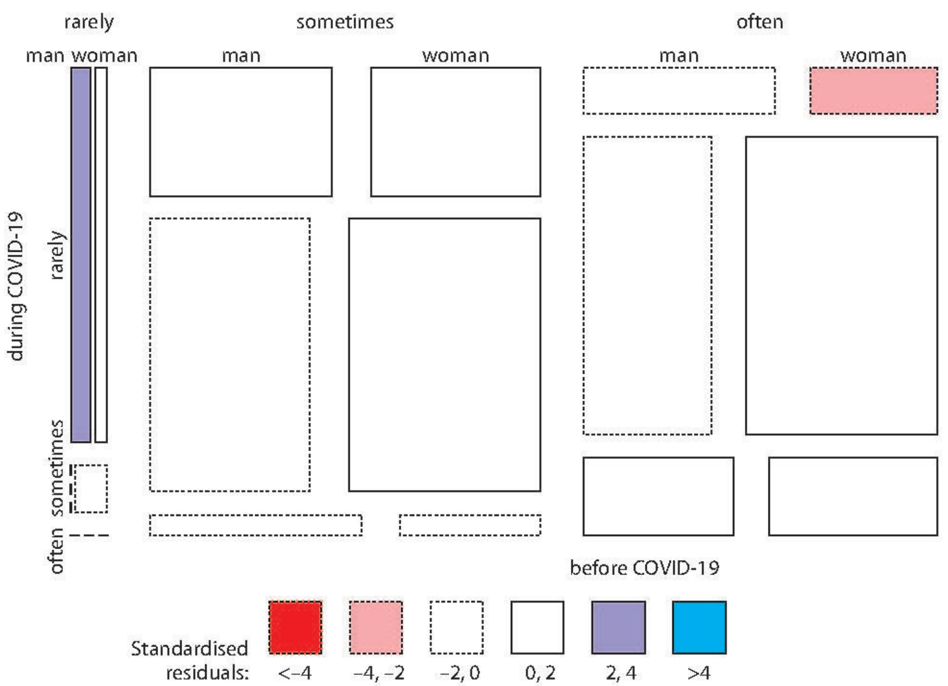
$$d_{ijk} = \frac{n_{ijk} - \hat{n}_{ijk}}{\sqrt{\hat{n}_{ijk}}}, \tag{4}$$

where $i$, $j$, $k$ are the numbers of rows, columns and layers.

The greater the standardised difference between the observed and expected frequencies, the darker the colour of the tile.

Two rectangles are distinguished in Figure 3. The red one represents the number of women who used to go to the cinema often before the pandemic, but rarely during COVID-19. The rectangle also indicates that there were fewer women than indicated by the expected frequencies. Assuming the independence of the three variables considered, there should be more women in this group. The blue rectangle, on the other hand, describes the number of men who rarely attended cinema shows before and during the COVID-19 pandemic. The observed numbers are larger than expected. Assuming the independence of the three variables considered, there should be fewer men. The coloured rectangles testify against the hypothesis of the independence of the three variables under study.

**Figure 3.** Participation in cinema screenings before and during the COVID-19 pandemic by gender




Note. Where the cell frequency is zero, the rectangle is reduced to a line, as indicated by the dashed lines.
Source: authors' work.

The constructed multidimensional contingency tables based on four sets of three variables have 18, 18, 36, and 24 cells, respectively, as shown in Table 5. In all of the considered three-dimensional contingency tables, there were cells for which the expected frequencies were below 5. In such cases, the chi-square independence test is not appropriate.

**Table 5.** Number of cells in multidimensional contingency tables and results of the chi-square independence test for four sets of three variables

| Variable | Number of cells | *p*-value[a] |
|---|---|---|
| Before.CV.19, During.CV.19, Gender | 18 | 2.63E-06 |
| Before.CV.19, During.CV.19, Age | 18 | 2.19E-05 |
| Before.CV.19, During.CV.19, Status | 36 | 2.97E-08 |
| Age, Status, Degree.of.compliance | 24 | 2.88E-09 |

[a] Chi-squared approximation may be incorrect.
Source: authors' work.

A chi-square test of independence was performed to check whether a statistically significant relationship between the three considered variables occurred. In each of the four sets of three variables, the *p*-value obtained in the R programme printouts was less than 0.001. In all cases, however, there was an indication that the chi-square approximation may not have been valid (Table 5). Since an inference based on the chi-square independence test was not possible, the study used a permutation test. The results of applying the permutation test are shown in Table 6.

**Table 6.** Permutation test results

| Variable | ASL ($N$ = 10,000) |
|---|---|
| Before.CV.19, During.CV.19, Gender | 0.0003 |
| Before.CV.19, During.CV.19, Age | 0.0001 |
| Before.CV.19, During.CV.19, Status | 0.0103 |
| Age, Status, Degree.of.compliance | 0.0002 |

Source: authors' work.

Each time, the ASL value is less than 0.05. In all of the analysed cases concerning the gender, age and status of the respondents, it was confirmed that these variables had an impact on the attendance at cinema screenings before and during the COVID-19 pandemic. The results were significantly influenced by the limited possibility to attend cinema screenings during the COVID-19 pandemic.

It was also confirmed that the Age and Status variables affected the degree to which respondents agreed with the statement: 'TV and online movies have replaced my ability to go to the cinema'.

## 6. Conclusions

Many issues, whether economic, social or medical, are in practice multidimensional. This article addresses the issue of statistical inference for multidimensional contingency tables. We use a permutation test based on Fisher's exact test for data in multidimensional contingency tables. When testing the independence of variables, the chi-square independence test is often used. If the contingency table has cells with expected frequencies smaller than 5, the chi-square approximation may prove inappropriate. If variables can take on multiple variants, then large sample sizes are required to infer independence. A permutation test allows statistical inference to be performed even when the multidimensional contingency table contains cells with expected frequencies below 5.

The paper provides examples of permutation tests used for examining data relating to participation in cinema screenings before and during the COVID-19 pandemic. The proposed method enabled the verification of the hypotheses of independence of the three variables, which was otherwise impossible using the chi-square independence test. The findings showed a statistically significant relationship between the variables included in the study.

## References

Agresti, A., & Coull, B. A. (1998). Approximate Is Better Than "Exact" for Interval Estimation of Binomial Proportions. *The American Statistician*, *52*(2), 119–126. https://doi.org/10.2307/2685469.

Agresti, A., & Gottard, A. (2007). Independence in multi-way contingency tables: S.N. Roy's breakthroughs and later developments. *Journal of Statistical Planning and Inference*, *137*(11), 3216–3226. https://doi.org/10.1016/j.jspi.2007.03.006.

Berger, R. L., & Boos, D. D. (1994). *P* Values Maximized Over a Confidence Set for the Nuisance Parameter. *Journal of the American Statistical Association*, *89*(427), 1012–1016. https://doi.org/10.1080/01621459.1994.10476836.

Berry, K. J., Johnston, J. E., & Mielke Jr., P. W. (2014). *A Chronicle of Permutation Statistical Methods: 1920–2000, and Beyond*. Springer. https://doi.org/10.1007/978-3-319-02744-9.

Berry, K. J., Johnston, J. E., & Mielke Jr., P. W. (2018). *The Measurement of Association. A Permutation Statistical Approach*. Springer. https://doi.org/10.1007/978-3-319-98926-6.

Berry, K. J., Mielke Jr., P. W., & Johnston, J. E. (2016). *Permutation Statistical Methods. An Integrated Approach*. Springer. https://doi.org/10.1007/978-3-319-28770-6.

Cheon, S.-Y. (2012). Approximating Exact Test of Mutual Independence in Multiway Contingency Tables via Stochastic Approximation Monte Carlo. *The Korean Journal of Applied Statistics*, *25*(5), 837–846. https://doi.org/10.5351/KJAS.2012.25.5.837.

Christensen, R. (1990). *Log-Linear Models*. Springer. https://doi.org/10.1007/978-1-4757-4111-7.

Domański, C. (1979). *Statystyczne testy nieparametryczne*. Państwowe Wydawnictwo Ekonomiczne.

Efron, B., & Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall. https://doi.org/10.1201/9780429246593.

Ernst, M. D. (2004). Permutation Methods: A Basis for Exact Inference. *Statistical Science*, *19*(4), 676–685. https://doi.org/10.1214/088342304000000396.

Fisher, R. A. (1935). *The Design of Experiments*. Hafner Press.

Fisz, M. (1967). *Rachunek prawdopodobieństwa i statystyka matematyczna*. Państwowe Wydawnictwo Naukowe.

Friendly, M. (1994). Mosaic Display for Multi-Way Contingency Tables. *Journal of the American Statistical Association*, *89*(425), 190–200. http://dx.doi.org/10.1080/01621459.1994.10476460.

Good, P. (2005). *Permutation, Parametric, and Bootstrap Tests of Hypotheses* (3rd edition). Springer. https://doi.org/10.1007/b138696.

Greń, J. (1974). *Statystyka matematyczna. Modele i zadania*. Państwowe Wydawnictwo Naukowe.

Kończak, G. (2012). On testing multi-directional hypotheses in categorical data analysis. In A. Colubi, K. Fokianos, E. J. Kontoghorges, K. Fokianos & G. Gonzalez-Rodrigues (Eds.), *20th International Conference on Computational Statistics* (pp. 427–436). International Statistical Institute.

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

Kończak, G., & Kosińska, M. (2023). O testowaniu istotności różnic w strukturach populacji na podstawie prób o małych liczebnościach. *Krakow Review of Economics and Management. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie*, (3), 145–160. https://doi.org/10.15678/ZNUEK.2023.1001.0308.

Kosińska, M. (2023). Wpływ pandemii COVID-19 na aktywność w wydarzeniach kulturalnych. *Studia Ekonomiczne. Gospodarka, Społeczeństwo, Środowisko. Economic Studies. Economy, Society, Environment*, (2), 61–73.

Little, R. J. A. (1989). Testing the Equality of Two Independent Binomial Proportions. *The American Statistician*, *43*(4), 283–288. https://doi.org/10.2307/2685390.

O'Gorman, T. W. (2012). *Adaptive Tests of Significance Using Permutations of Residuals with R and SAS.* Wiley.

Perrone, E., Fontana, R., & Rapallo, F. (2024). *Multi-way contingency tables with uniform margins.* http://arxiv.org/abs/2404.04343.

Raevskikh, E., Khalid, U., & Benghozi, P. J. (2022). *Culture in Times of COVID-19. Resilience, Recovery and Revival.* UNESCO. https://www.nck.pl/upload/2023/08/culture-in-times-of-covid-19-resilience-recovery-and-revival.pdf.

Rao, C. R. (1973). *Linear Statistical Inference and its Applications* (2nd edition). John Wiley & Sons. https://doi.org/10.1002/9780470316436.

Santander. (2023). *Sytuacja społeczno-gospodarcza Polski w dobie pandemii*. https://odpowiedzialnybiznes.pl/wp-content/uploads/2020/12/FOB_Santander_Sytuacja_spoleczno-gospodarcza_Polski_w_dobie_pandemii.pdf.

Sheskin, D. J. (2004). *Handbook of parametric and nonparametric statistical procedures* (3rd edition). Chapman & Hall/CRC.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Science*. McGraw-Hill.

Simonoff, J. S. (2003). *Analyzing Categorical Data*. Springer. https://doi.org/10.1007/978-0-387-21727-7.

Stanimir, A. (2011). Analysis of Nominal Data – Multi-way Contingency Table. *Mathematical Economics*, *7*(14), 229–240.

Yates, F. (1934). Contingency Tables Involving Small Numbers and the $\chi^2$ Test. *Supplement to the Journal of the Royal Statistical Society*, *1*(2), 217–235. https://doi.org/10.2307/2983604.

Zeliaś, A., Pawełek, B., & Wanat, S. (2002). *Metody statystyczne. Zadania i sprawdziany*. Polskie Wydawnictwo Ekonomiczne.