

Piotr SULEWSKI¹

Nonparametric versus parametric reasoning based on two-way and three-way contingency tables

1. INTRODUCTION

The analysis of contingency tables (CTs) is one of the most common tasks performed by statisticians. CTs display the frequency distribution of two (two-way CTs), three (three-way CTs) or more (multi-way CTs) categorical variables. The information about categorical data can be found e.g. in Bishop et al. (1975), Agresti (2002), Van Belle et al. (2014). Information presented as CTs features in a wide variety of areas such as the social sciences (Wickens, 1969), genetics (El Galta et al., 2008; Dickhaus et al., 2012), demography (Cung, 2013) and psychology (Iossifova et al., 2013). Basic methods of testing for dependency in CTs in details is described e.g. in Steinle et al. (2006), Bock (2003), Kaski et al. (2005), Allison, Liker (1982). Other examples of applications may be found in Ilyas et al. (2004), Oates, Cohen (1996), Schrepp (2003), Haas et al. (2007).

One can recognize two general cases in which CTs can be useful. This distinction between the cases is made with respect to the tasks which CTs are used for.

Case A. Dependency is unwanted. The general population is sought to be in its normal state or be under control when levels of feature X are independent of levels of feature Y . Revealing dependency means revealing abnormality of members of the general population. If so, a large scale and very costly actions have to be obligatorily initiated. That is why a decision-maker tries to avoid false alarm. This case is typical, for instance, in security guarding. A classic statistical way of reasoning is tailored to case A. Please notice that the main hypothesis, commonly denoted by H_0 , states that: X and Y are independent. Moreover, H_0 is guarded against rejection by setting significance level at 5% or less.

Case B. Dependency is wanted. The state of the general population is assessed upon feature X . Unfortunately, levels of feature X are difficult to be determined e.g. determination is risky, costly or time consuming. In contrast, levels of another feature Y are easy to be determined. Assessors are concerned with finding out whether there is a tie between X and Y . In other words, whether X

¹ The Pomeranian University, Institute of Mathematics, 6–7 Koziętułskiego St., 76–200 Słupsk, Poland, e-mail: piotr.sulewski@apsl.edu.pl.

and Y are dependent or independent. Assessors use Y_1, \dots, Y_k levels as sensible indicators of X_1, \dots, X_w levels. Case B is typical in diagnostics, both medical and technical. In case B another way of statistical reasoning is needed, different from the classic way.

Conservativeness of the classic statistical way of reasoning often obstructs progress in numerous situations where rejecting H_0 means making a step ahead. This is a strong motivation for making a turnaround in statistical reasoning. In this new statistical reasoning there is no null hypothesis. In contrast to the classic way, there is a set of competing hypotheses. Moreover, the testing procedure warrants equality of all the alternatives when the test begins. The former null hypothesis is no longer the main one, but exists among the other ones of equal importance. Particular hypotheses relate to scenarios under which particular CTs are created. Details are presented in section 7. There are two reasons for which this likelihood based reasoning is developed and put forward:

- a) Undoubtedly, CT-based classic statistical reasoning is the nonparametric reasoning. It is commonly known that parametric statistical reasoning, if applicable, is much more sensitive to untruthfulness of H_0 than nonparametric reasoning. In this paper we propose a parametric reasoning. Particular scenarios are parameterized with the probability flow parameter (PFP).
- b) Let us again retrace a way of the classic thinking. A value of the test statistics is smaller than the appropriate critical value results in failing to reject H_0 . In case A the decision maker is comfortable about independence. A value of the test statistics is no smaller than the appropriately determined critical value results in rejecting H_0 . In case B the decision maker is comfortable about dependence because there is no word said what the reason of rejecting H_0 is. The most likely scenario is selected whereas reasons to reject H_0 or not are embedded in scenarios. One can say that the method put forward in this paper offers a transition from "unfathomable" to "fathomable" reasons.

Nonparametric and parametric reasoning based on 2×2 CT is presented in (Sulewski, 2018b), therefore this paper is devoted to bigger tables, e.g. 2×3 , 2×4 , 3×3 , 3×4 , 4×4 , $2 \times 2 \times 2$, $3 \times 2 \times 2$ ones.

This paper is organized as follows. Variants of presentation of CT are described in section 2. CT coming into being are presented in section 3. Statistic tests including the power divergence tests and the $|\chi|$ test are defined in section 4. Section 5 is devoted to measures of untruthfulness of H_0 including the measure that is defined by means of an absolute value. In section 6 maximum likelihood method is applied to estimate the PFP. Section 7 is devoted to instructions how to generate two-way and three-way CTs. Section 8 presents numerical examples and section 9 presents closing remarks.

Monte Carlo simulation is performed in Visual Basic for Applications embedded in Microsoft Excel 2016.

2. VARIANTS OF PRESENTATION OF CT

This section is devoted to the $w \times k \times p$ CT. If $p = 1$, then we obviously have $w \times k$ CT.

Let X, Y, Z be three features of the same object, respectively, have levels $X_1, \dots, X_w, Y_1, \dots, Y_k, Z_1, \dots, Z_p$. Testing these three features for independence with an appropriately arranged CT is probably one of the most common statisticians' tasks. At the moment one can distinct between four variants of presentation of CTs. It is because each variant is intended for a different purpose. Below details of particular variants are given:

- TP Variant (theoretical probabilities). Cells contain probabilities p_{ijt} intrinsic to the phenomenon being investigated (see table 1). The exact values of these probabilities are unknown to the investigator. This variant is introduced a little bit in advance since CTs will be simulated with the Monte Carlo method in further sections of this paper. And just then CT variant filled with probabilities arbitrarily set by Monte Carlo experimenter will be applied.
- TC Variant (theoretical counts). Cells contain theoretical expected counts $n_{ijt} = np_{ijt}$. These counts are theoretical in this sense that they result from TP variant.
- EP Variant (empirical probabilities). Cells that result from EC variant and contain estimates $p_{ijt}^* = n_{ijt}^*/n$ of the unknown content of TP.
- EC Variant (experimental counts). Cells contain n_{ijt}^* counts observed on a sample drawn from general population subjected to the investigation.

Table 1. TP VARIANT OF THREE-WAY CT

Z	Z ₁			...	Z _p			Total
X/Y	Y ₁	...	Y _k	...	Y ₁	...	Y _k	
Y ₁	P_{111}	...	P_{1k1}	...	P_{11p}	...	P_{1kp}	$P_{1..}$
...
X _w	P_{w11}	...	P_{wk1}	...	P_{w1p}	...	P_{wkp}	$P_{w..}$
Total	$P_{.11}$...	$P_{.k1}$...	$P_{.1p}$...	$P_{.kp}$	1

3. ON HOW CT COMES INTO BEING

One can treat CTs as a mathematical expression of a certain phenomenon we deal with. This formulation suggests that there is an internal mechanism in this phenomenon that determines probabilities of particular X, Y or X, Y, Z combinations and ascribes these probabilities to the cells of the table. Below are "progenitors" of all the $w \times k$ CTs

$$T_p = \begin{bmatrix} 1/(wk) & \dots & 1/(wk) \\ \vdots & \dots & \vdots \\ 1/(wk) & \dots & 1/(wk) \end{bmatrix} \tag{1}$$

and all the $w \times k \times p$ CTs

$$W_p = \begin{bmatrix} 1/(wkp) & \dots & 1/(wkp) & \dots & 1/(wkp) & \dots & 1/(wkp) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1/(wkp) & \dots & 1/(wkp) & \dots & 1/(wkp) & \dots & 1/(wkp) \end{bmatrix} \tag{2}$$

A variety of tables may be generated when portions of PFP a flow from "maternal" cells of (1) or (2) to other cells. Obviously, the total probability always equals 1. In this paper twenty eight scenarios that seem fundamental are developed (tables 2–3). These scenarios are created based on scenarios for the 2×2 CT that seem fundamental and describe different levels of dependence (Sulewski, 2018b).

Table 2. THE CONTENTS OF $w \times k$ CTs RESULTING FROM SCENARIOS IN QUESTION

Table	Scenario	Content
2 × 3	I	$p_{11} = 1/6 - a, p_{12} = p_{13} = p_{21} = p_{22} = 1/6, p_{23} = 1/6 + a$
	II	$p_{11} = p_{12} = 1/6 - a, p_{13} = p_{21} = 1/6, p_{22} = p_{23} = 1/6 + a$
	III	$p_{11} = p_{12} = 1/6 - a, p_{13} = p_{23} = 1/6, p_{21} = p_{22} = 1/6 + a$
	IV	$p_{11} = p_{23} = 1/6 - a, p_{12} = p_{22} = 1/6, p_{13} = p_{31} = 1/6 + a$
2 × 4	V	$p_{11} = 1/8 - a, p_{12} = p_{13} = p_{14} = p_{21} = p_{22} = p_{23} = 1/8, p_{24} = 1/8 + a$
	VI	$p_{11} = p_{12} = p_{13} = 1/8 - a, p_{14} = p_{21} = 1/8, p_{22} = p_{23} = p_{24} = 1/8 + a$
	VII	$p_{11} = 1/8 - a, p_{12} = p_{13} = p_{14} = p_{22} = p_{23} = p_{24} = 1/8, p_{21} = 1/8 + a$
	VIII	$p_{11} = p_{12} = 1/8 - a, p_{13} = p_{14} = p_{23} = p_{24} = 1/8, p_{21} = p_{22} = 1/8 + a$
3 × 3	IX	$p_{11} = 1/9 - a, p_{12} = p_{13} = p_{21} = p_{22} = p_{23} = p_{31} = p_{32} = 1/9, p_{33} = 1/9 + a$
	X	$p_{11} = p_{12} = 1/9 - a, p_{13} = p_{21} = p_{22} = p_{23} = p_{33} = 1/9, p_{31} = p_{32} = 1/9 + a$
	XI	$p_{11} = p_{33} = 1/9 - a, p_{12} = p_{21} = p_{22} = p_{23} = p_{32} = 1/9, p_{13} = p_{31} = 1/9 + a$
	XII	$p_{11} = p_{21} = p_{33} = 1/9 - a, p_{12} = p_{22} = p_{32} = 1/9, p_{13} = p_{23} = p_{31} = 1/9 + a$
3 × 4	XIII	$p_{11} = 1/12 - a, p_{12} = p_{13} = p_{14} = p_{21} = p_{22} = p_{23} = p_{24} = p_{31} = p_{32} = p_{33} = 1/12, p_{34} = 1/12 + a$
	XIV	$p_{11} = p_{12} = 1/12 - a, p_{13} = p_{14} = p_{21} = p_{22} = p_{23} = p_{24} = p_{33} = p_{34} = 1/12, p_{31} = p_{32} = 1/12 + a$
	XV	$p_{11} = p_{12} = 1/12 - a, p_{21} = p_{22} = p_{23} = p_{24} = p_{31} = p_{32} = p_{33} = p_{34} = 1/12, p_{13} = p_{14} = 1/12 + a$
	XVI	$p_{11} = p_{12} = p_{33} = p_{34} = 1/12 - a, p_{21} = p_{22} = p_{23} = p_{24} = 1/12, p_{13} = p_{14} = p_{31} = p_{32} = 1/12 + a$

Table 2. THE CONTENTS OF $w \times k$ CTs RESULTING FROM SCENARIOS IN QUESTION (cont.)

Table	Scenario	Content
4 × 4	XVII	$p_{11} = p_{12} = p_{13} = 1/16 - a, p_{14} = p_{21} = p_{22} = p_{23} = p_{24} = p_{31} = p_{32} = p_{33} =$ $= p_{34} = p_{44} = 1/16, p_{41} = p_{42} = p_{43} = 1/16 + a$
	XVIII	$p_{11} = p_{12} = 1/16 - a, p_{13} = p_{14} = p_{21} = p_{22} = p_{23} = p_{24} = p_{31} = p_{32} = p_{33} =$ $= p_{34} = p_{41} = p_{42} = 1/16, p_{43} = p_{44} = 1/16 + a$
	XIX	$p_{11} = p_{12} = p_{13} = p_{21} = p_{22} = p_{23} = 1/16 - a, p_{14} = p_{24} = p_{21} =$ $= p_{31} = p_{41} = 1/16, p_{32} = p_{33} = p_{34} = p_{42} = p_{43} = p_{44} = 1/16 + a$
	XX	$p_{11} = p_{12} = p_{43} = p_{44} = 1/16 - a, p_{21} = p_{22} = p_{23} = p_{24} = p_{31} =$ $p_{32} = p_{33} = p_{34} = 1/16, p_{13} = p_{14} = p_{41} = p_{42} = 1/16 + a$

Source: own elaboration.

Table 3. THE CONTENTS OF $w \times k \times p$ CTs RESULTING FROM SCENARIOS IN QUESTION

Table	Scenario	Content
2 × 2 × 2	XXI	$p_{111} = p_{112} = 1/8 - a, p_{121} = p_{122} = p_{211} = p_{212} = 1/8, p_{221} = p_{222} = 1/8 + a$
	XXII	$p_{111} = p_{121} = p_{112} = 1/8 - a, p_{122} = p_{211} = 1/8, p_{221} = p_{212} = p_{222} = 1/8 + a$
	XXIII	$p_{111} = 1/8 - a, p_{121} = p_{112} = p_{122} = p_{221} = p_{212} = p_{222} = 1/8, p_{211} = 1/8 + a$
	XXIV	$p_{111} = p_{112} = 1/8 - a, p_{121} = p_{122} = p_{221} = p_{222} = 1/8, p_{211} = p_{212} = 1/8 + a$
3 × 2 × 2	XXV	$p_{111} = p_{112} = 1/12 - a, p_{121} = p_{122} = p_{211} = p_{221} = p_{212} = p_{222} = p_{311} = p_{312} =$ $= 1/12, p_{321} = p_{322} = 1/12 + a$
	XXVI	$p_{111} = p_{112} = 1/12 - a, p_{121} = p_{211} = p_{221} = p_{212} = p_{222} = p_{311} =$ $= p_{312} = p_{322} = 1/12, p_{122} = p_{321} = 1/12 + a$
	XXVII	$p_{111} = p_{112} = p_{211} = 1/12 - a, p_{212} = p_{222} = p_{311} = p_{321} =$ $= p_{312} = p_{322} = 1/12, p_{121} = p_{122} = p_{221} = 1/12 + a$
	XXVIII	$p_{111} = p_{112} = p_{321} = p_{322} = 1/12 - a, p_{211} = p_{221} = p_{212} =$ $= p_{222} = 1/12, p_{121} = p_{122} = p_{311} = p_{312} = 1/12 + a$

Source: own elaboration.

In all the above scenarios the PFP a takes values in $\left[0, \frac{1}{wp}\right]$ or $\left[0, \frac{1}{wpk}\right]$. The scenarios are selected in such a way that they correspond to different levels of dependence expressed by means of an appropriate measure of untruthfulness of H_0 (MoU). The MoU takes values on interval $[0,1]$. A simulation study is carried out for MoU values no bigger than $2/3$. It is obvious that the detection of a strong dependence is very simple. You can find more information about the MoU in section 5.

Obviously, scenarios do not cover all the cases. They may be locally mutated by reversing rows or columns to better fit the analyzed data. These are simple equal-portion scenarios. In the scenarios you can use a part of PFP, e.g. $a/2$, $a/3, \dots$. Surely, real scenarios can be more or less similar to these above. This is typical in relations between theory and real life. With the current availability of computers, the statistician can afford situations that interest him and instantly repeat such simulations. All examples presented here have a very precise algorithmic description in a form of a step list.

The researches can be generalized by introducing several PFPs. This, however, causes a significant deterioration in the properties of the parameter estimators. The Weibull distribution has a simple analytical form. For its generalization, the Generalized Gamma Distribution (URG) can be considered. Due to big problems with estimating URG parameters the author does not know any practical applications of URG to describe the reliability results of technical objects. You can always add more parameters to the model, however, this might worsen their estimation.

4. INDEPENDENCE TESTS

4.1. Two-way contingency table

Features X, Y are independent what means that H_0 is true, if $p_{ij} = p_{i\cdot} \cdot p_{\cdot j}(i, j = 1, 2)$ for each pair of i, j . The alternative hypothesis denoted H_1 is such one that negates H_0 . Let e_{ij} be the expected counts

$$e_{ij} = \frac{n_{i\cdot} \cdot n_{\cdot j}}{n} = np_{i\cdot} \cdot p_{\cdot j}(i = 1, \dots, w; j = 1, \dots, k). \tag{3}$$

The expected counts e_{ij} have the same one-way marginal values as the observed table n_{ij} (Gokhale, Kullback, 1978).

Statistical science has been enriched with many other statistics intended for research on test independency. Cressie, Read (1984) propose the power divergence statistics (PDS). The PDS for $w \times k$ CTs is given by

$$P^2 = \frac{2}{\lambda(\lambda + 1)} \sum_{i=1}^w \sum_{j=1}^k n_{ij}^* \left[\left(\frac{n_{ij}^*}{e_{ij}^*} \right)^\lambda - 1 \right] = \tag{4}$$

$$= \frac{2n}{\lambda(\lambda + 1)} \sum_{i=1}^w \sum_{j=1}^k p_{ij}^* \left[\left(\frac{p_{ij}^*}{p_{i\cdot}^* \cdot p_{\cdot j}^*} \right)^\lambda - 1 \right] - \infty < \lambda < \infty,$$

where e_{ij} values are given by (3). Equation (4) always takes positive values and is defined as a limit of P^2 at -1 and 0 . P^2 contains a very rich class of test statistics, for example: the χ^2 statistics ($\lambda = 1$), the G^2 statistics (the limit as λ goes to 0), the Freeman-Tukey statistics ($\lambda = -0.5$), the modified G^2 statistics (the limit as λ goes to -1), the Neyman modified χ^2 statistics ($\lambda = -2$) and the Cressie-Read statistics ($\lambda = 2/3$). If H_0 is true, statistics (4), for large n (i.e. asymptotically), follows the chi-square distribution with $(w - 1)(k - 1)$ degrees of freedom.

The following PDS are selected to Monte Carlo study: the χ^2 statistics (Pearson, 1900), the Freeman-Tukey FT statistics (Freeman, Tukey, 1950), the Cressie-Read CR statistics (Cressie, Read, 1984):

$$\chi^2 = \sum_{i=1}^w \sum_{j=1}^k \frac{(n_{ij}^* - e_{ij})^2}{e_{ij}}, \quad (5)$$

$$FT = 4 \sum_{i=1}^w \sum_{j=1}^k \left(\sqrt{n_{ij}^*} - \sqrt{e_{ij}^*} \right)^2, \quad (6)$$

$$CR = \frac{9}{5} \sum_{i=1}^w \sum_{j=1}^k n_{ij}^* \left[\left(\frac{n_{ij}^*}{e_{ij}^*} \right)^{2/3} - 1 \right]. \quad (7)$$

The G^2 statistics (Sokal, Rohlf, 2012), the modified G^2 statistics (Kullback, 1959) and the Neyman modified χ^2 statistics (Neyman, 1949) have not been subjects in the Monte Carlo study because they are applicable only in a case where all n_{ij} ($i = 1, \dots, w; j = 1, \dots, k$) counts are not equal to zero.

The square used in the numerator of χ^2 statistics (5) makes that large differences between expected and theoretical counts even bigger and the small differences even smaller. Another aim of the use of the square is to avoid that the differences are mutually exclusive. For this purpose one can use their absolute value instead of squared deviations. The $|\chi|$ statistics is selected to Monte Carlo study, too. It is an authorial modification of χ^2 statistics and it has the form (Sulewski, 2013)

$$|\chi| = \sum_{i=1}^w \sum_{j=1}^k \frac{|n_{ij}^* - e_{ij}^*|}{e_{ij}^*} = \sum_{i=1}^w \sum_{j=1}^k \frac{|p_{ij}^* - p_{i\cdot}^* p_{\cdot j}^*|}{p_{i\cdot}^* p_{\cdot j}^*}, \quad (8)$$

where n_{ij}^* are experimental counts, e_{ij}^* are expected counts and p_{ij}^* are empirical probabilities. It is shown in (Sulewski, 2016) that $|\chi|$ test is more powerful than tests (5)–(7).

4.2. Three-way contingency table

In this paper the research has been limited only to complete independence. Features X, Y, Z are completely independent from one another, and H_0 is true, if

$$P_{ijt} = P_{i..}P_{.j.}P_{..t} \tag{9}$$

for each $i = 1, \dots, w; j = 1, \dots, k; t = 1, \dots, p$. The alternative hypothesis H_1 negates H_0 .

Let e_{ijt} be the expected counts under complete independence of i, j, t

$$e_{ijt} = \frac{n_{i..}n_{.j.}n_{..t}}{n^2} = np_{i..}p_{.j.}p_{..t} (i = 1, \dots, w; j = 1, \dots, k; t = 1, \dots, p). \tag{10}$$

The expected counts e_{ijt} under complete independence of i, j, t have the same one-way

marginal values as the observed table n_{ijt} (Gokhale, Kullback, 1978)

To study the complete independence of the features X, Y, Z we use the statistics that are extensions of those for two-way CTs (Pardo, 1996)

$$\chi_3^2 = \sum_{i=1}^w \sum_{j=1}^k \sum_{t=1}^p \frac{(n_{ijt}^* - e_{ijt}^*)^2}{e_{ijt}^*}, \tag{11}$$

$$FT_3 = 4 \sum_{i=1}^w \sum_{j=1}^k \sum_{t=1}^p \left(\sqrt{n_{ijt}^*} - \sqrt{e_{ijt}^*} \right)^2, \tag{12}$$

$$CR_3 = \frac{9}{5} \sum_{i=1}^w \sum_{j=1}^k \sum_{t=1}^p n_{ijt}^* \left[\left(\frac{n_{ijt}^*}{e_{ijt}^*} \right)^{2/3} - 1 \right]. \tag{13}$$

Statistics (11)–(13) for CT, when H_0 is true, asymptotically follow the chi-square distribution with $wkp - (w + k + p) + 2$ degrees of freedom. Statistics

$$G_3^2 = 2 \sum_{i=1}^w \sum_{j=1}^k \sum_{t=1}^p n_{ijt}^* \ln \left(\frac{n_{ijt}^*}{e_{ijt}^*} \right), N_3 = \sum_{i=1}^w \sum_{j=1}^k \sum_{t=1}^p \frac{(n_{ijt}^* - e_{ijt}^*)^2}{n_{ijt}^*},$$

$$KL_3 = 2 \sum_{i=1}^w \sum_{j=1}^k \sum_{t=1}^p e_{ijt}^* \ln \left(\frac{e_{ijt}^*}{n_{ijt}^*} \right)$$

also belong to the PDS. However, these statistics have not been applied in the Monte Carlo study, because they do not take into account the condition $n_{ijt}^* = 0$.

The $|\chi_3|$ statistics is selected to Monte Carlo study, too. It is an authorial modification of χ_3^2 statistics and it has the form (Sulewski, 2018a)

$$|\chi_3| = \sum_{i=1}^w \sum_{j=1}^k \sum_{t=1}^p \frac{|n_{ijt}^* - e_{ijt}^*|}{e_{ijt}^*}, \quad (14)$$

where n_{ijt}^* are experimental counts and e_{ijt}^* are expected counts. It is shown in (Sulewski, 2018a) that $|\chi_3|$ test is more powerful than tests (11)–(13).

5. MEASURES OF UNTRUTHFULNESS OF H_0

5.1. Two-way contingency table

When the equality $p_{ij} = p_{i\cdot}p_{\cdot j}$ is not fulfilled, H_0 is not true and an appropriate measure of untruthfulness of H_0 (MoUH) is needed. There are many different measures in literature, e.g.: the Pearson's ϕ , the Tschuprow's T , the Cramer's V , the corrected contingency c , the Goodman and Kruskal's τ .

In this paper we use a MoUH which is given by (Sulewski, 2016):

$$MoU = \frac{1}{n} \sum_{i=1}^w \sum_{j=1}^k \left| n_{ij}^* - \frac{n_{i\cdot}^* \cdot n_{\cdot j}^*}{n} \right| = \sum_{i=1}^w \sum_{j=1}^k |p_{ij}^* - p_{i\cdot}^* \cdot p_{\cdot j}^*|. \quad (15)$$

The MoU takes values in interval $\langle 0,1 \rangle$. This measure, doubtlessly, springs from the essence of H_0 and has a very simple form. The MoU formulas and the maximal MoU values (the minimal MoU values are equal to zero) under scenarios I-XX are presented in table 4. The MoU is a function of the PFP a . Owing to this the MoU values are very easy to calculate.

Table 4. THE MoU UNDER SCENARIOS I-XX FOR A TWO-WAY CT

Table	Scenario	MoU	MoU_{max}	Table	Scenario	MoU	MoU_{max}
2 × 3	I	$4a/3$	0.2222	2 × 4	V	$2a$	0.25
	II	*	0.3333		VI	**	0.3125
	III	$8a/3$	0.4444		VII	$3a$	0.375
	IV	$4a$	0.6667		VIII	$4a$	0.5
3 × 3	IX	$2a^2 + 2a$	0.2469	3 × 4	XIII	$8a/3$	0.2222
	X	$8a/3$	0.2963		XIV	$4a$	0.3333
	XI	$4a$	0.4444		XV	$16a/3$	0.4444
	XII	$16a/3$	0.5926		XVI	$8a$	0.6667

Table 4. THE MoU UNDER SCENARIOS I-XX FOR A TWO-WAY CT (cont.)

Table	Scenario	MoU	MoU_{max}	Table	Scenario	MoU	MoU_{max}
4×4	XVII	$3a$	0.1875				
	XVIII	$4a$	0.25				
	XIX	***	0.3125				
	XX	$8a$	0.5				

* $4a/3 I(0 \leq a < 0.1), 8a^2 + 2a/3 I(0.1 < a \leq 1/6)$
 ** $2a I(0 \leq a \leq 0.075), 12a^2 + a I(0.075 < a \leq 0.125)$
 *** $4a I(0 \leq a \leq 0.0375), 48a^2 + 2a I(0.0375 < a \leq 1/16)$

Source: own elaboration.

5.2. Three-way contingency table

The theory devoted to $MoUH$ for TT is not as rich as for the two-way contingency table, where the Goodman—Kruskal index plays an important role (Goodman, Kruskal, 1954). Numeric extensions of this index for three way CT are: the Marcotorchino index τ_M (Marcotorchino, 1984), the delta index τ_L (Lombardo, 2011) and the Gray—Williams index τ_{GW} (Gray, Williams, 1975). Information about other less popular indices can be found in (Beh, Davy, 1998; Harshman, 1970; Lombardo, Beh, 2010; Trucker, 1963).

Based on the classical definition of independence of X, Y, Z , the MoU_3 in the form

$$MoU_3 = \frac{1}{n} \sum_{i=1}^w \sum_{j=1}^k \sum_{t=1}^p \left| n_{ijt}^* - \frac{n_{i..}^* \cdot n_{.j.}^* \cdot n_{..t}^*}{n^2} \right| = \sum_{i=1}^w \sum_{j=1}^k \sum_{t=1}^p |p_{ijt} - p_{i..} \cdot p_{.j.} \cdot p_{..t}|, \quad (16)$$

is put forward in (Sulewski, 2018a). The measure (16) takes the value 0 when H_0 is true. The higher the MoU_3 value, the greater the possibility of H_0 falsity. More information about the measures MoU_3, τ_M, τ_L and τ_{GW} defined under some scenarios for three-way CT you can find in (Sulewski, 2018a).

The MoU_3 as a natural measure, resulting from the definition of independence, is used in the Monte Carlo simulation. The MoU_3 formulas and the maximal MoU_3 values (the minimal MoU_3 values are equal to zero) under scenarios XXI–XXVIII are presented in table 5. The MoU_3 is a function of the PFP a . Owing to this the MoU_3 values are very easy to calculate.

Table 5. The MoU_3 under scenarios XXI-XXVIII for three-way contingency tables

Table	Scenario	MoU_3	$MoU_{3\ max}$	Table	Scenario	MoU_3	$MoU_{3\ max}$
$2 \times 2 \times 2$	XXI	$16a^2$	0.25	$3 \times 2 \times 2$	XXV	$8a/3$	0.2222
	XXII	*	0.3359		XXVI	**	0.3889
	XXIII	$3a$	0.375		XXVII	$6a$	0.5
	XXIV	$4a$	0.5		XXVIII	$8a$	0.6667

*

<i>a</i>	0	0.0125	0.025	0.0375	0.05	0.0625	0.075	0.0875	0.1	0.1125	0.125
<i>MoU₃</i>	0	0.025	0.0504	0.0763	0.103	0.1309	0.1601	0.1969	0.238	0.283	0.3359

**

<i>a</i>	0	0.0083	0.0167	0.025	0.0333	0.0417	0.05	0.0583	0.0667	0.075	0.0833
<i>MoU₃</i>	0	0.0364	0.0733	0.1108	0.1489	0.1875	0.2267	0.2664	0.3067	0.3475	0.3889

Source: own elaboration.

6. APPLYING THE MAXIMUM LIKELIHOOD METHOD TO ESTIMATE THE PROBABILITY FLOW PARAMETER

This section is a simply attempt of replacing a nonparametric statistical inference method by the parametric one. Maximum likelihood method is applied to estimate the PFP.

Let us remember that in cells of two-way CT are values n_{ij} ($i = 1, \dots, w; j = 1, \dots, k$), in cells of three-way CT are values n_{ijt} ($i = 1, \dots, w; j = 1, \dots, k; t = 1, \dots, p$). These values are components of the multinomial distribution. Thus the multinomial distribution was taken as a groundwork of the likelihood functions. A family of these likelihood functions is given below. Every function from this family has an index. Indices assign functions to particular scenarios presented in section 3 of the paper.

6.1. Two-way contingency table

Let n_{ij}^* be the value of (i, j) cell and a is the probability flow parameter. Then likelihood functions under CT $w \times k$ have the form

a) table 2×3

$$L_I(a) = C(1/6 - a)^{n_{11}^*} (1/6 + a)^{n_{23}^*} (1/6)^{n_{12}^* + n_{13}^* + n_{21}^* + n_{22}^*}, \quad (17)$$

$$L_{II}(a) = C(1/6 - a)^{n_{11}^* + n_{12}^*} (1/6 + a)^{n_{22}^* + n_{23}^*} (1/6)^{n_{13}^* + n_{21}^*}, \quad (18)$$

$$L_{III}(a) = C(1/6 - a)^{n_{11}^* + n_{12}^*} (1/6 + a)^{n_{22}^* + n_{23}^*} (1/6)^{n_{13}^* + n_{21}^*}, \quad (19)$$

$$L_{IV}(a) = C(1/6 - a)^{n_{11}^* + n_{23}^*} (1/6 + a)^{n_{13}^* + n_{21}^*} (1/6)^{n_{12}^* + n_{22}^*}, \quad (20)$$

b) table 2×4

$$L_V(a) = C(1/8 - a)^{n_{11}^*} (1/8 + a)^{n_{24}^*} (1/8)^{n_{12}^* + n_{13}^* + n_{14}^* + n_{21}^* + n_{22}^* + n_{23}^*}, \quad (21)$$

$$L_{VI}(a) = C(1/8 - a)^{n_{11}^* + n_{12}^* + n_{13}^*} (1/8 + a)^{n_{22}^* + n_{23}^* + n_{24}^*} (1/8)^{n_{14}^* + n_{21}^*}, \quad (22)$$

$$L_{VII}(a) = C(1/8 - a)^{n_{12}^*} (1/8 + a)^{n_{21}^*} (1/8)^{n - n_{11}^* - n_{21}^*}, \quad (23)$$

$$L_{VIII}(a) = C(1/8 - a)^{n_{11}^* + n_{12}^*} (1/8 + a)^{n_{21}^* + n_{22}^*} (1/8)^{n_{13}^* + n_{14}^* + n_{23}^* + n_{24}^*}, \quad (24)$$

c) table 3×3

$$L_{IX}(a) = C(1/9 - a)^{n_{11}^*}(1/9 + a)^{n_{33}^*}(1/9)^{n - n_{11}^* - n_{33}^*}, \quad (25)$$

$$L_X(a) = C(1/9 - a)^{n_{11}^* + n_{12}^*}(1/9 + a)^{n_{31}^* + n_{31}^*}(1/9)^{n_{13}^* + n_{21}^* + n_{22}^* + n_{23}^* + n_{33}^*}, \quad (26)$$

$$L_{XI}(a) = C(1/9 - a)^{n_{11}^* + n_{33}^*}(1/9 + a)^{n_{11}^* + n_{33}^*}(1/9)^{n_{12}^* + n_{21}^* + n_{22}^* + n_{23}^* + n_{32}^*}, \quad (27)$$

$$L_{XII}(a) = C(1/9 - a)^{n_{11}^* + n_{21}^* + n_{33}^*}(1/9 + a)^{n_{11}^* + n_{23}^* + n_{33}^*}(1/9)^{n_{12}^* + n_{22}^* + n_{32}^*}, \quad (28)$$

d) table 3×4

$$L_{XIII}(a) = C(1/12 - a)^{n_{11}^*}(1/12 + a)^{n_{34}^*}(1/12)^{n - n_{11}^* - n_{34}^*}, \quad (29)$$

$$L_{XIV}(a) = C(1/12 - a)^{n_{11}^* + n_{12}^*}(1/12 + a)^{n_{31}^* + n_{32}^*}(1/12)^{n - n_{11}^* - n_{12}^* - n_{31}^* - n_{32}^*}, \quad (30)$$

$$L_{XV}(a) = C(1/12 - a)^{n_{11}^* + n_{12}^*}(1/12 + a)^{n_{13}^* + n_{14}^*}(1/12)^{n - n_{11}^* - n_{12}^* - n_{13}^* - n_{14}^*}, \quad (31)$$

$$L_{XVI}(a) = C(1/12 - a)^{n_{11}^* + n_{33}^* + n_{34}^*}(1/12 + a)^{n_{13}^* + n_{14}^* + n_{31}^* + n_{32}^*} \\ (1/12)^{n_{21}^* + n_{22}^* + n_{23}^* + n_{24}^*}, \quad (32)$$

e) table 4×4

$$L_{XVII}(a) = C(1/16 - a)^{n_{11}^* + n_{12}^* + n_{13}^*}(1/16 + a)^{n_{41}^* + n_{42}^* + n_{43}^*} \\ (1/16)^{n - n_{11}^* - n_{12}^* - n_{13}^* - n_{41}^* - n_{42}^* - n_{43}^*}, \quad (33)$$

$$L_{XVIII}(a) = C(1/16 - a)^{n_{11}^* + n_{12}^*}(1/16 + a)^{n_{43}^* + n_{44}^*}(1/16)^{n - n_{11}^* - n_{12}^* - n_{43}^* - n_{44}^*}, \quad (34)$$

$$L_{XIX}(a) = C(1/16 - a)^{n_{11}^* + n_{12}^* + n_{13}^* + n_{21}^* + n_{22}^* + n_{23}^*} \\ (1/16 + a)^{n_{32}^* + n_{33}^* + n_{34}^* + n_{42}^* + n_{43}^* + n_{44}^*}(1/16)^{n - n_{14}^* - n_{24}^* - n_{31}^* - n_{41}^*}, \quad (35)$$

$$L_{XX}(a) = C(1/16 - a)^{n_{11}^* + n_{12}^* + n_{43}^* + n_{44}^*} \\ (1/16 + a)^{n_{13}^* + n_{14}^* + n_{41}^* + n_{42}^*}(1/16)^{n_{21}^* + n_{22}^* + n_{23}^* + n_{24}^* + n_{31}^* + n_{32}^* + n_{33}^* + n_{34}^*}, \quad (36)$$

In formulas (17)–(36) $C = n! / \prod_{i=1}^w \prod_{j=1}^k n_{ij}^*$.

The logarithmic likelihood function under scenario I in CT 2×3 is given by

$$l_I(a) = 1nL_I(a) = 1n(C) + n_{11}^* 1n(1/6 + a) + (n_{12}^* + n_{13}^* n_{21}^* n_{22}^*) 1n(1/6),$$

then

$$\frac{\partial l_I(a)}{\partial a^2} = \frac{n_{23}^*}{1/6 + a} - \frac{\partial l_I(a)}{\partial a} = 0 \Rightarrow \hat{a} = \frac{n_{23}^* - n_{11}^*}{6(n_{23}^* + n_{11}^*)}$$

Let us check what kind of extremum we can find. As a result of a simple transformation we have

$$\frac{\partial l_I(a)}{\partial a^2} = \frac{\partial}{\partial a} \left[\frac{n_{23}^*}{1/6 + a} - \frac{n_{11}^*}{1/6 - a} \right] = -\frac{n_{23}^*}{(1/6 + a)^2} - \frac{n_{11}^*}{(1/6 - a)^2} < 0, \quad (37)$$

for $a < 1/6$. It means that the logarithmic likelihood function has always a maximum at $a = \hat{a}$. So, \hat{a} is the maximum likelihood estimator of a which is the probability flow parameter. It may be proven that inequality (37) holds for all scenarios considered in this paper.

Formulas for the maximum likelihood estimator of a under scenarios I-XX for two-way CT are given by:

a) table 2×3

$$\begin{aligned} \hat{a}_I &= \frac{n_{23}^* - n_{11}^*}{6(n_{23}^* + n_{11}^*)}, \hat{a}_{II} = \frac{n_{22}^* + n_{23}^* - (n_{11}^* + n_{12}^*)}{6(n_{11}^* + n_{12}^* + n_{22}^* + n_{23}^*)}, \\ \hat{a}_{III} &= \frac{n_{21}^* + n_{22}^* - (n_{11}^* + n_{12}^*)}{6(n_{11}^* + n_{12}^* + n_{21}^* + n_{22}^*)}, \hat{a}_{IV} = \frac{n_{13}^* + n_{21}^* - (n_{11}^* + n_{23}^*)}{6(n_{11}^* + n_{12}^* + n_{21}^* + n_{23}^*)}, \end{aligned} \quad (38)$$

b) table 2×4

$$\begin{aligned} \hat{a}_V &= \frac{n_{24}^* - n_{11}^*}{8(n_{24}^* + n_{11}^*)}, \hat{a}_{VI} = \frac{n_{22}^* + n_{23}^* + n_{24}^* - (n_{11}^* + n_{12}^* + n_{13}^*)}{8(n_{11}^* + n_{12}^* + n_{21}^* + n_{22}^* + n_{23}^* + n_{24}^*)}, \\ \hat{a}_{VII} &= \frac{n_{21}^* - n_{11}^*}{8(n_{11}^* + n_{21}^*)}, \hat{a}_{VIII} = \frac{n_{21}^* + n_{22}^* - (n_{11}^* + n_{12}^*)}{8(n_{11}^* + n_{12}^* + n_{21}^* + n_{22}^*)}, \end{aligned} \quad (39)$$

c) table 3×3

$$\begin{aligned} \hat{a}_{XI} &= \frac{n_{33}^* - n_{11}^*}{9(n_{11}^* + n_{33}^*)}, \hat{a}_X = \frac{n_{31}^* + n_{32}^* - (n_{11}^* + n_{12}^*)}{9(n_{11}^* + n_{12}^* + n_{31}^* + n_{32}^*)}, \\ \hat{a}_{XII} &= \frac{n_{13}^* + n_{31}^* - (n_{11}^* + n_{33}^*)}{9(n_{11}^* + n_{13}^* + n_{31}^* + n_{33}^*)}, \hat{a}_{XIII} = \frac{n_{13}^* + n_{23}^* + n_{31}^* - (n_{11}^* + n_{21}^* + n_{33}^*)}{9(n_{11}^* + n_{12}^* + n_{22}^* + n_{32}^*)}, \end{aligned} \quad (40)$$

d) table 3×4

$$\begin{aligned} \hat{a}_{XIII} &= \frac{n_{34}^* - n_{11}^*}{12(n_{11}^* + n_{34}^*)}, \hat{a}_{XIV} = \frac{n_{31}^* + n_{32}^* - (n_{11}^* + n_{12}^*)}{12(n_{11}^* + n_{12}^* + n_{31}^* + n_{32}^*)}, \\ \hat{a}_{XV} &= \frac{n_{13}^* + n_{14}^* - (n_{11}^* + n_{12}^*)}{12(n_{11}^* + n_{12}^* + n_{13}^* + n_{14}^*)}, \\ \hat{a}_{XVI} &= \frac{n_{13}^* + n_{14}^* + n_{31}^* + n_{32}^* - (n_{11}^* + n_{12}^* + n_{33}^* + n_{34}^*)}{12(n_{11}^* + n_{12}^* + n_{22}^* + n_{23}^* + n_{24}^*)}, \end{aligned} \quad (41)$$

e) table 4×4

$$\begin{aligned} \hat{a}_{XVII} &= \frac{n_{41}^* + n_{42}^* + n_{43}^* - (n_{11}^* + n_{12}^* + n_{13}^*)}{16(n_{11}^* + n_{12}^* + n_{13}^* + n_{41}^* + n_{42}^* + n_{43}^*)}, \\ \hat{a}_{XVIII} &= \frac{n_{43}^* + n_{44}^* - (n_{11}^* + n_{12}^*)}{16(n_{11}^* + n_{12}^* + n_{43}^* + n_{44}^*)}, \\ \hat{a}_{XIX} &= \frac{n_{32}^* + n_{33}^* + n_{34}^* + n_{42}^* + n_{43}^* + n_{44}^* - (n_{11}^* + n_{12}^* + n_{13}^* + n_{21}^* + n_{22}^* + n_{23}^*)}{16(n - n_{14}^* + n_{24}^* + n_{31}^* + n_{41}^*)}, \\ \hat{a}_{XX} &= \frac{n_{13}^* + n_{14}^* + n_{41}^* + n_{42}^* - (n_{11}^* + n_{12}^* + n_{31}^* + n_{44}^*)}{16(n_{11}^* + n_{12}^* + n_{13}^* + n_{14}^* + n_{41}^* + n_{42}^* + n_{43}^* + n_{44}^*)}. \end{aligned} \tag{42}$$

To decide which of the defined scenarios takes place, you use the following algorithm:

1. Find dimension of CT in question and read out related q according to the rule: $2 \times 3 (q = 1), 2 \times 4 (q = 2), 3 \times 3 (q = 3), 3 \times 4 (q = 4), 4 \times 4 (q = 5)$
2. Set a set of scenario indices $\{4q - 3, 4q - 2, 4q - 1, 4q\}$.
3. Calculate a^* , which is an estimate of parameter a for each scenario from step 2 by means of (38)–(42).
4. Calculate corresponding values of the maximum likelihood functions $L_{4q-3}(a^*), L_{4q-2}(a^*), L_{4q-1}(a^*), L_{4q}(a^*)$ by means of (17)–(36).
5. Choose a scenario for which the value $L(a^*)$ is the greatest.

This algorithm will be used in section 8.1, Example 1.

6.2. Three-way contingency table

Let n_{ijt}^* be the value of (i, j, t) cell and a is the PFP. Then likelihood functions for selected three-way CT have the form:

a) table $2 \times 2 \times 2$

$$L_{XXI}(a) = D(1/8 - a)^{n_{111}^* + n_{112}^*} (1/8 + a)^{n_{221}^* + n_{222}^*} (1/8)^{n_{221}^* + n_{211}^* n_{122}^* + n_{212}^*}, \tag{43}$$

$$L_{XXII}(a) = D(1/8 - a)^{n_{111}^* + n_{121}^* n_{122}^*} (1/8 + a)^{n_{221}^* + n_{212}^* + n_{222}^*} (1/8)^{n_{122}^* + n_{211}^*}, \tag{44}$$

$$L_{XXIII}(a) = D(1/8 - a)^{n_{111}^*} (1/8 + a)^{n_{211}^*} (1/8)^{n - n_{111}^* + n_{211}^*}, \tag{45}$$

$$L_{XXIV}(a) = D(1/8 - a)^{n_{111}^* + n_{112}^*} (1/8 + a)^{n_{211}^* + n_{212}^*} (1/8)^{n_{121}^* + n_{221}^* + n_{122}^* + n_{222}^*}, \tag{46}$$

b) table $3 \times 2 \times 2$

$$L_{XXV}(a) = D(1/8 - a)^{n_{111}^* + n_{112}^*} (1/8 + a)^{n_{321}^* + n_{322}^*} (1/8)^{n - n_{111}^* - n_{112}^* - n_{321}^* - n_{322}^*}, \tag{47}$$

$$L_{XXVI}(a) = D(1/8 - a)^{n_{111}^* + n_{112}^*} (1/8 + a)^{n_{122}^* + n_{321}^*} (1/8)^{n - n_{111}^* - n_{112}^* - n_{122}^* - n_{321}^*}, \tag{48}$$

$$\begin{aligned} L_{XXVII}(a) &= D(1/8 - a)^{n_{111}^* + n_{112}^* + n_{221}^*} (1/8 + a)^{n_{121}^* + n_{122}^* + n_{211}^*} \\ &\quad (1/8)^{n_{212}^* + n_{222}^* + n_{311}^* + n_{321}^* + n_{312}^* + n_{222}^* + n_{322}^*}, \end{aligned} \tag{49}$$

$$L_{XXVIII}(a) = D(1/8 - a)^{n_{111}^* + n_{112}^* + n_{321}^* + n_{322}^*} (1/8 + a)^{n_{121}^* + n_{122}^* + n_{311}^* + n_{312}^*} (1/8)^{n_{211}^* + n_{221}^* + n_{212}^* + n_{221}^*}. \quad (50)$$

In formulas (43)–(50) $D = n! / \prod_{i=1}^w \prod_{j=1}^k \prod_{t=1}^p n_{ijt}^*$.

Formulas for the maximum likelihood estimator of a under scenarios XXIV–XXVIII for CT $w \times k \times p$ are given by

a) table $2 \times 2 \times 2$

$$\begin{aligned} \hat{a}_{XXI} &= \frac{n_{221}^* + n_{222}^* - (n_{111}^* + n_{112}^*)}{8(n_{111}^* + n_{112}^* + n_{221}^* + n_{222}^*)}, \\ \hat{a}_{XXII} &= \frac{n_{221}^* + n_{212}^* + n_{222}^* - (n_{111}^* + n_{121}^* + n_{122}^*)}{8(n - n_{122}^* - n_{211}^*)}, \\ \hat{a}_{XXIII} &= \frac{n_{211}^* - n_{111}^*}{8(n_{111}^* + n_{211}^*)}, \hat{a}_{XXIV} = \frac{n_{211}^* + n_{212}^* - (n_{111}^* + n_{112}^*)}{8(n_{111}^* + n_{112}^* + n_{211}^* + n_{212}^*)}. \end{aligned} \quad (51)$$

a) table $3 \times 2 \times 2$

$$\begin{aligned} \hat{a}_{XXV} &= \frac{n_{321}^* + n_{322}^* - (n_{111}^* + n_{112}^*)}{12(n_{111}^* + n_{112}^* + n_{321}^* + n_{322}^*)}, \\ \hat{a}_{XXVI} &= \frac{n_{122}^* + n_{321}^* - (n_{111}^* + n_{112}^*)}{12(n_{111}^* + n_{112}^* + n_{122}^* + n_{321}^*)}, \\ \hat{a}_{XXVII} &= \frac{n_{121}^* + n_{122}^* + n_{211}^* - (n_{111}^* + n_{112}^* + n_{221}^*)}{8(n_{111}^* + n_{112}^* + n_{211}^* + n_{212}^*)}, \\ \hat{a}_{XXVIII} &= \frac{n_{121}^* + n_{112}^* + n_{311}^* + n_{312}^* - (n_{111}^* + n_{112}^* + n_{321}^* + n_{322}^*)}{12(n - n_{211}^* - n_{221}^* - n_{212}^* - n_{222}^*)}. \end{aligned} \quad (52)$$

To decide which of the defined scenarios takes place, you use the following algorithm:

1. Find dimension of CT in question and read out related q according to the rule:

$$2 \times 2 \times 2 (q = 6), 3 \times 2 \times 2 (q = 7)$$

2. Set a set of scenario indices $\{4q - 3, 4q - 2, 4q - 1, 4q\}$.
3. Calculate a^* , which is an estimate of parameter a for each scenario from step 2 according to (51)–(52)
4. Calculate corresponding values of the maximum likelihood functions $L_{4q-3}(a^*), L_{4q-2}(a^*), L_{4q-1}(a^*), L_{4q}(a^*)$ by means of (43)–(50).
5. Choose a scenario for which the value $L(a^*)$ is the greatest.

This algorithm will be used in section 8.2, Example 4.

7. GENERATING CONTINGENCY TABLE

Generating CT is very important in the simulation study. The approach in the literature for the generating two-way CTs is the Markov Chain Monte Carlo (Diaconis, Sturmfels, 1998; Cryan, Dyer, 2003; Cryan et al., 2006; Chen et al., 2005; Fishman, 2012), the Sequential Importance Sampling (Chen et al., 2005; Chen et al., 2006; Blitzstein, Diaconis, 2011; Yoshida, 2011), the probabilistic

divide-and-conquer technique (DeSalvo, Zhao, 2015), the Generalized Gamma Distribution (Sulewski, 2009), the bar method (Sulewski, Motyka, 2015). The bar method for the generating three-way CT is presented in (Sulewski, 2018a).

In this paper we use an algorithm for generating two-way and three-way CTs using the bar method. The bar method is identical to the method that generates random numbers that follow the multinomial distribution.

8. PARAMETRIC REASONING PUT INTO PRACTICE

8.1. Two-way contingency table

Example 1. This example compares decision making by means of classic statistical testing with likelihood based decision making. Tables 6-10 show a set of example two-way CTs filled one by one accordingly to the scenarios I–XX. The table is divided into two parts. The left hand side is related to likelihood based decisions. To decide which of the defined scenarios takes place, see algorithm in section 6.1. The right hand side is related to classic statistical testing and presents values of test statistics. The H_0 states that X and Y are independent. Critical values, indicated by underlining, are determined by Monte Carlo method based on 10^6 order statistics. Such a large number of repetitions guaranties very precise results. When reading rows of the table it turns out that all the decisions made in a classic way are wrong. It is because untrue H_0 hypotheses have not been rejected. But it does not reveal anything new. This is just one more confirmation of what is commonly known – the classic statistical test is very conservative. The PFP α is a maximal value of this parameter for which untrue H_0 is rejected in no scenarios.

Tables 6–10 show that all the decisions made in a classic way are wrong under the scenarios in question. It is because untrue H_0 hypotheses have not been rejected even in situations where the MoU does not have such small values, e.g. $MoU = 0.3$ in scenario XX. In turn, the parametric approach detected a dependency between features.

Example 2. The conservativeness of classic testing is a reason why we suggest making a turnaround in this domain. The new proposal is a method of statistical inference and not a classical parametric test. Now there will be no null hypothesis, but there will be a set of competing alternative hypotheses instead. The former H_0 is no longer the main one, but exists among the competitors of an equal rank. All the hypotheses state: “the considered two-way CT is generated accordingly to particular scenarios”. Each figure from 1 to 5 shows sets of four likelihood curves. Each curve has its four attributes, namely \hat{a} , n , table dimensions and, of course, the actual generation scenario that is specified in the figure's title. The legend lists all competing scenarios including the actual one.

Table 6. THE CLASSIC STATISTICAL TESTING VERSUS LIKELIHOOD BASED DECISIONS, TABLE 2 x 3

Experiment settings: $n = 60, \alpha = 0.05, \alpha = 0.05$

Scenario (MoU)	Content		X and Y dependent	Scenario of maximum likelihood	χ^2	FT	CR	X	H_0 true	H_0 rejected
I(0.067)	7	10	Yes	I	6.041	6.422	6.083	1.767	No	No
	10	10								
II(0.067)	7	7	Yes	II	0.334	0.338	0.335	0.402	No	No
	10	13								
III(0.133)	7	7	Yes	III	1.250	1.237	1.246	0.833	No	No
	13	13								
IV(0.200)	7	10	Yes	IV	3.600	3.706	3.612	1.200	No	No
	13	7								

Source: own elaboration.

Table 7. THE CLASSIC STATISTICAL TESTING VERSUS LIKELIHOOD BASED DECISIONS, TABLE 2 x 4

Experiment settings: $n = 80, \alpha = 0.0625, \alpha = 0.05$

Scenario (MoU)	Content				X and Y dependent	Scenario of maximum likelihood	χ^2	FT	CR	X	H_0 true	H_0 rejected
	5	10	10	10								
V(0.125)	5	10	10	10	Yes	V	1.439	1.462	1.443	1.084	No	No
	10	10	10	15								
VI(0.125)	5	5	5	10	Yes	VI	1.648	1.642	1.645	1.086	No	No
	10	15	15	15								
VII(0.188)	5	10	10	10	Yes	VII	3.810	4.104	3.855	1.524	No	No
	15	10	10	10								
VIII(0.250)	5	5	10	10	Yes	VIII	5.333	5.482	5.350	2.133	No	No
	15	15	10	10								

Source: own elaboration.

Table 8. THE CLASSIC STATISTICAL TESTING VERSUS LIKELIHOOD BASED DECISIONS, TABLE 3 × 3

Experiment settings: $n = 90, \alpha = 0.0444, \alpha = 0.05$

Scenario (MoU)	Content			X and Y dependent	Scenario of maximum likelihood	χ^2	FT	CR	x	H_0 true	H_0 rejected
	6	10	10								
IX(0.093)	6	10	10	Yes	XI	1.138	1.160	1.142	0.870	No	No
	10	10	10								
	10	10	14								
X(0.119)	6	6	10	Yes	X	2.297	2.257	2.283	1.148	No	No
	10	10	10								
	14	14	10								
XI(0.178)	6	10	14	Yes	XI	6.400	6.750	6.440	1.600	No	No
	10	10	10								
	14	10	6								
XII(0.237)	6	10	14	Yes	XII	8.688	8.991	8.690	2.172	No	No
	6	10	14								
	14	10	6								

Source: own elaboration.

Table 9. THE CLASSIC STATISTICAL TESTING VERSUS LIKELIHOOD BASED DECISIONS, TABLE 3 x 4

Experiment settings: $n = 80, \alpha = 0.0625, \alpha = 0.05$

Scenario (MoU)	Content			X and Y dependent	Scenario of maximum likelihood	χ^2	FT	CR	X	H_0 true	H_0 rejected
XIII(0.067)	7	10	10	Yes	XIII	0.770	0.781	0.772	0.813	No	No
	10	10	10								
	10	10	10								
XIV(0.100)	7	7	10	Yes	XIV	1.841	1.856	1.843	1.228	No	No
	10	10	10								
	13	13	10								
XV(0.133)	7	7	13	Yes	XV	2.424	2.474	2.433	1.616	No	No
	10	10	10								
	10	10	10								
XVI(0.200)	7	7	13	Yes	XVI	7.200	7.413	7.225	2.400	No	No
	10	10	10								
	13	13	7								

Source: own elaboration.

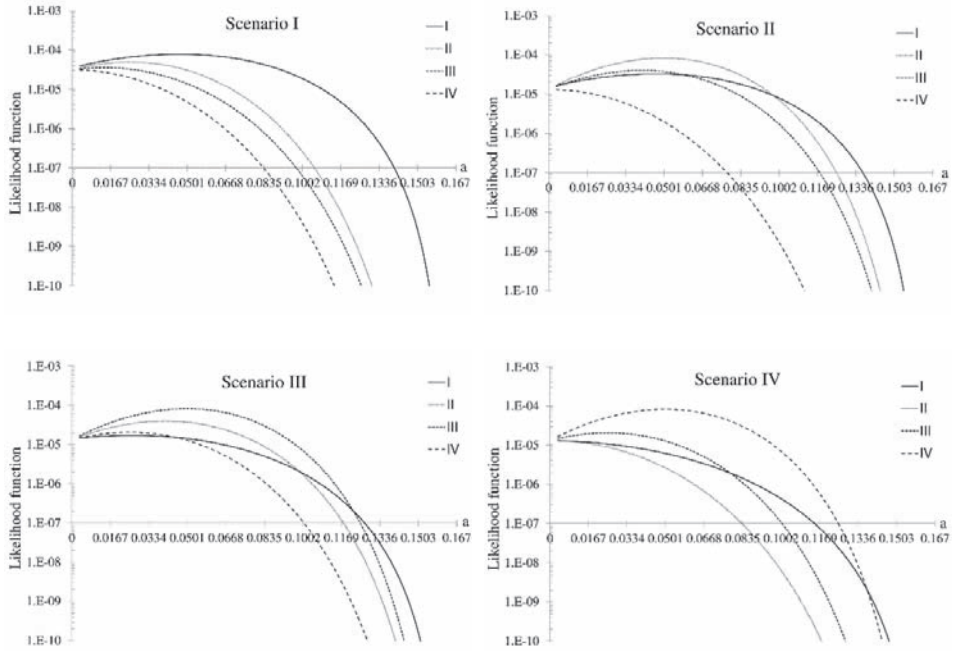
Table 10. THE CLASSIC STATISTICAL TESTING VERSUS LIKELIHOOD BASED DECISIONS, TABLE 4 × 4

Experiment settings: $n = 80, \alpha = 0.0375, \alpha = 0.05$

Scenario (MoU)	Content				X and Y dependent	Scenario of maximum likelihood	χ^2	FT	CR	X	H_0 true	H_0 rejected
	2	2	2	5								
XVII(0.113)	2	2	2	5	Yes	XVII	3.386	3.105	3.297	2.257	No	No
	5	5	5	5								
	5	5	5	5								
	8	8	8	5								
XVIII(0.150)	2	2	5	5	Yes	XVIII	2.206	2.277	2.217	2.698	No	No
	5	5	5	5								
	5	5	5	5								
	5	5	8	8								
XIX(0.150)	2	2	2	5	Yes	XIX	2.704	2.677	2.690	2.712	No	No
	2	2	2	5								
	5	8	8	8								
	5	8	8	8								
XX(0.300)	2	2	8	8	Yes	XX	14.400	16.421	14.613	4.800	No	No
	5	5	5	5								
	5	5	5	5								
	8	8	2	2								

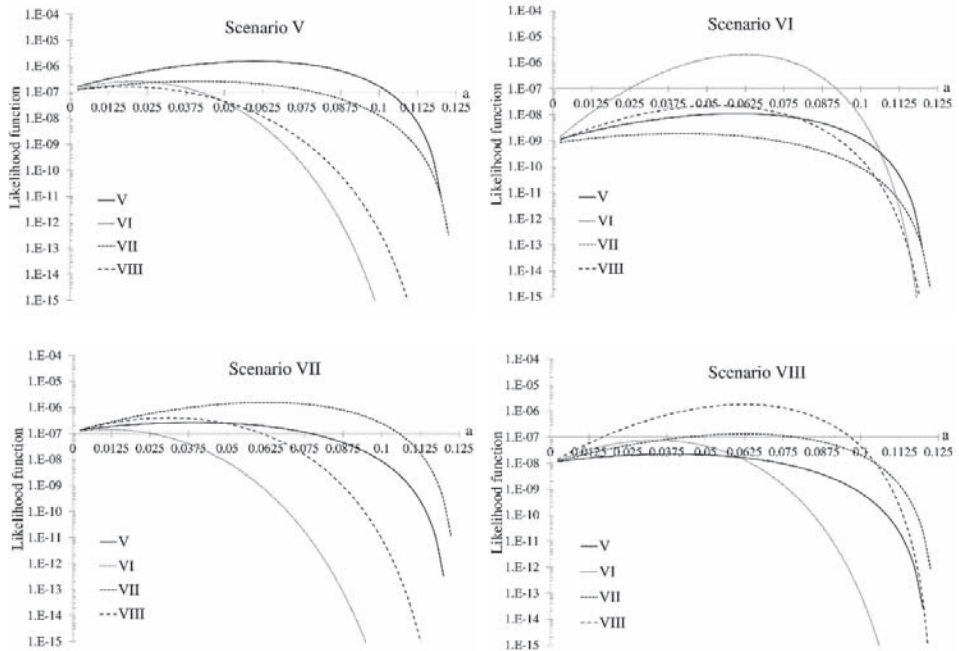
Source: own elaboration.

Figure 1. The likelihood functions for $\hat{a} = 0.05, n = 60$, table 2×3

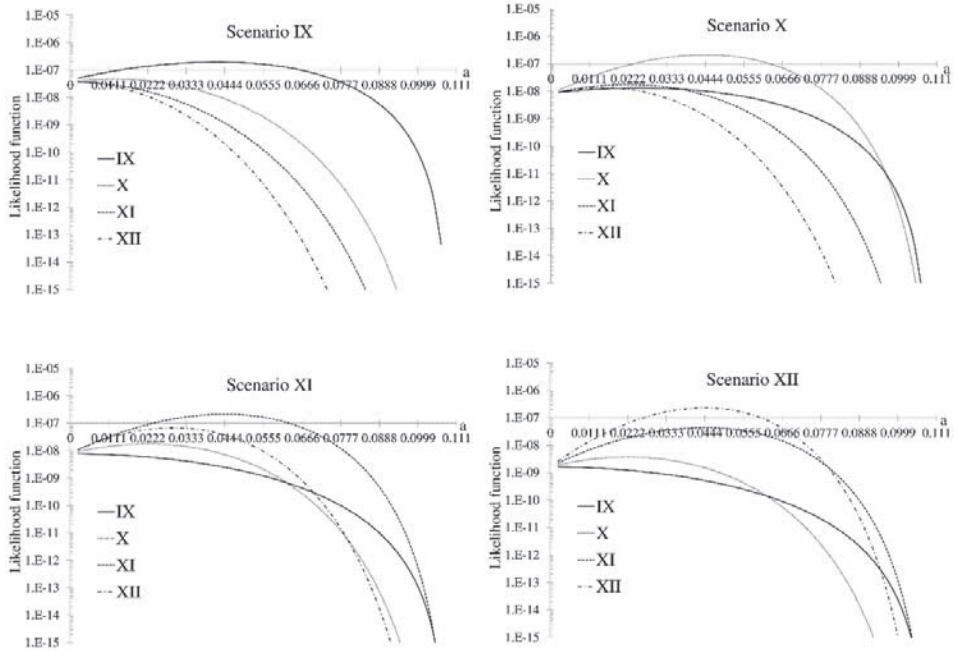


Source: own elaboration.

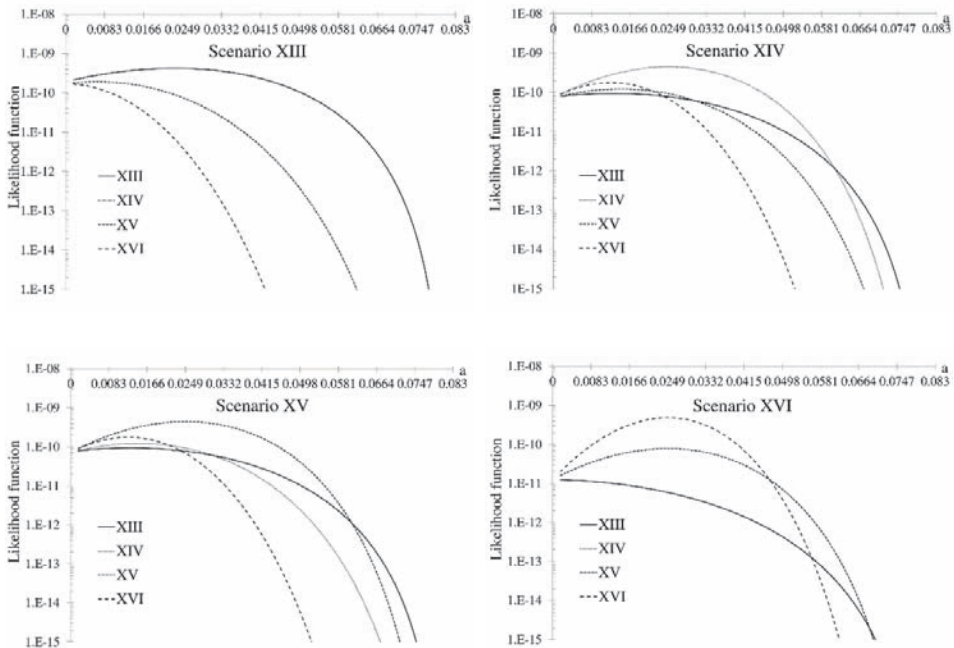
Figure 2. The likelihood functions for $\hat{a} = 0.0625, n = 80$, table 2×4



Source: own elaboration.

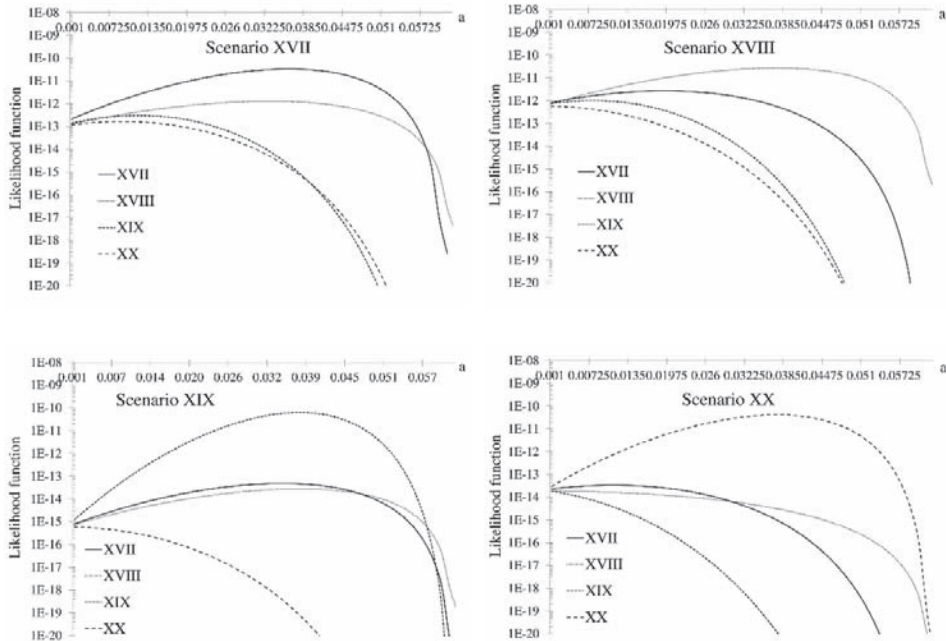
Figure 3. The likelihood functions for $\hat{a} = 0.0444, n = 90$, table 3×3 

Source: own elaboration.

Figure 4. The likelihood functions for $\hat{a} = 0.0250, n = 120$, table 3×4 

Source: own elaboration.

Figure 5. The likelihood functions for $\hat{a} = 0.0375, n = 80$, table 4×4



Source: own elaboration.

It is noteworthy (see figures 1–5) that in each set of likelihood curves, the curve related to the actual scenario predominates over the others. It is instructive to read out a value of a_{max} that maximizes likelihood function on a particular figure and notice that this value is close to assumed \hat{a} .

Example 3. This example is carried out in accordance with the following algorithm:

1. Find dimension of CT in question and read out related q according to the rule:

$$2 \times 3 (q = 1), 2 \times 4 (q = 2), 3 \times 3 (q = 3), 3 \times 4 (q = 4), 4 \times 4 (q = 5)$$

2. Set a set of scenario indices $\{4q - 3, 4q - 2, 4q - 1, 4q\}$.
3. Calculate values of PFP by means of $a_i = 0.1i/(wk)$ ($i = 1, 2, \dots, 10$).
4. Set a sample size n .
5. Repeat the following steps $u = 10^4$ times:
 - 5.1. Let $Sc_{4q-3} = 0, Sc_{4q-2} = 0, Sc_{4q-1} = 0, Sc_{4q} = 0$
 - 5.2. Generate CT under the scenarios that you have chosen in Step 2.
 - 5.3. Calculate $L_{4q-3}(a), L_{4q-2}(a), L_{4q-1}(a), L_{4q}(a)$ by means of (17)–(36).
 - 5.4. If $MaxL = L_{4q-3}(a)$, then $Sc_{4q-3} = Sc_{4q-3} + 1$,
 If $MaxL = L_{4q-2}(a)$, then $Sc_{4q-2} = Sc_{4q-2} + 1$,

If $MaxL = L_{4q-1}(a)$, then $Sc_{4q-1} = Sc_{4q-1} + 1$,

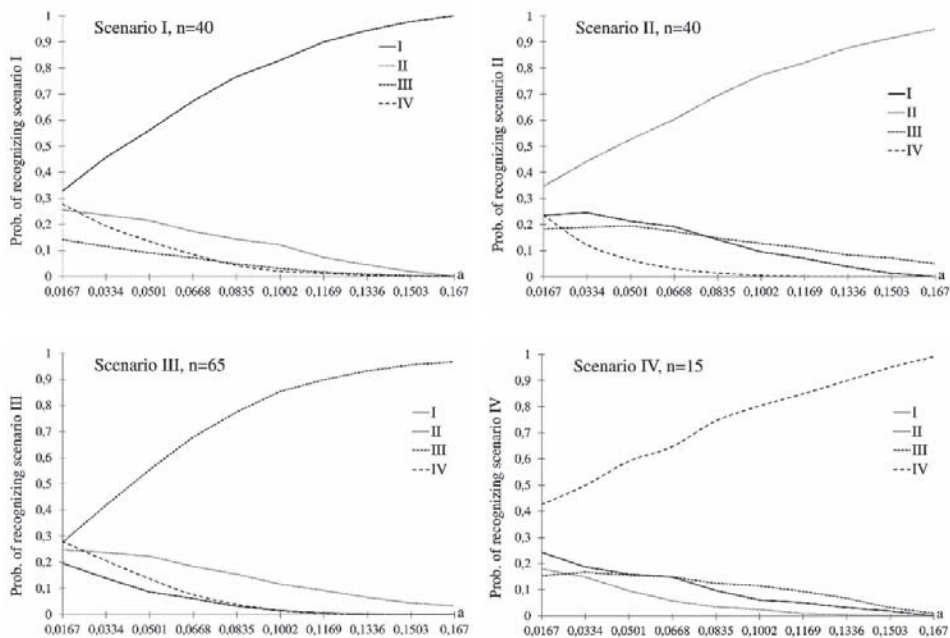
If $MaxL = L_{4q}(a)$, then $Sc_{4q} = Sc_{4q} + 1$.

6. Calculate probabilities of recognizing (PoR) scenarios. These are probabilities for the actual scenario to be recognized as one of scenarios in question.

$$Pr_{4q-3} = Sc_{4q-3}/u, Pr_{4q-2} = Sc_{4q-2}/u, Pr_{4q-1} = Sc_{4q-1}/u, Pr_{4q} = Sc_{4q}/u.$$

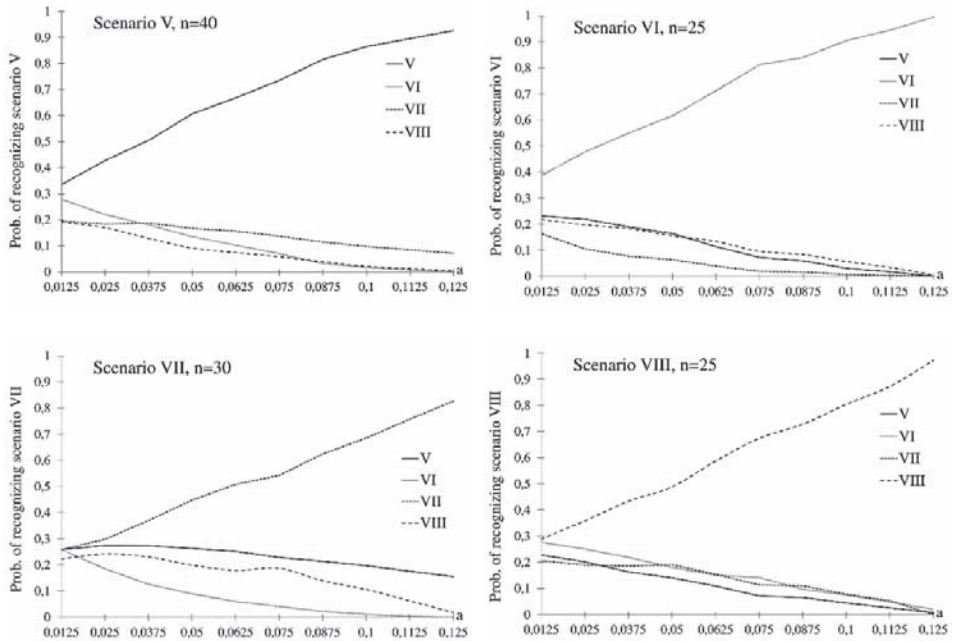
Figures 6-10 present PoR(a) functions for two-way CT in question. Sample sizes for a given CT are different because maximal MoU values under the scenarios are different (see table 4). The minimal sample sizes are chosen in such a way that probabilities of proper recognition (actual I as I, ..., actual XX as XX) are greater than probabilities of improper recognitions for all the a values.

Figure 6. The PoR actual scenario as one of I–IV scenarios, table 2×3



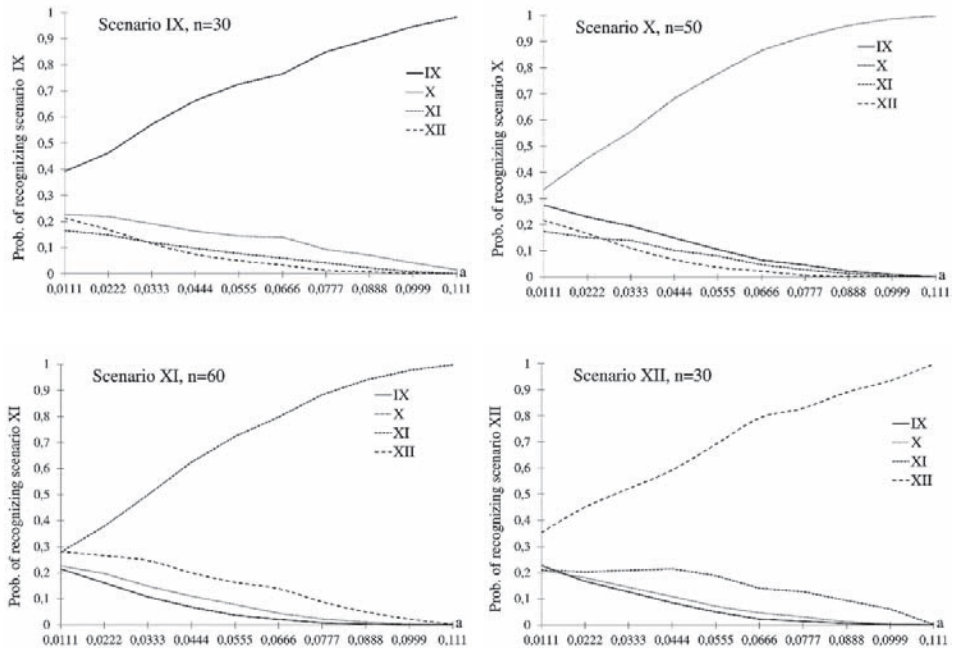
Source: own elaboration.

Figure 7. The PoR actual scenario as one of V–VIII scenarios, table 2 × 4



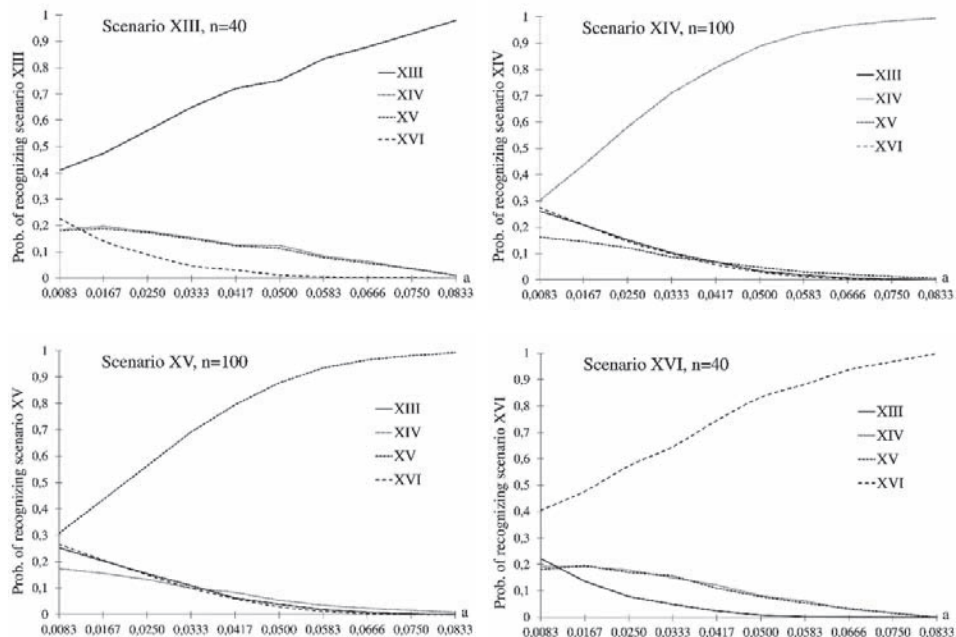
Source: own elaboration.

Figure 8. The PoR actual scenario as one of IX–XII scenarios, table 3 × 3



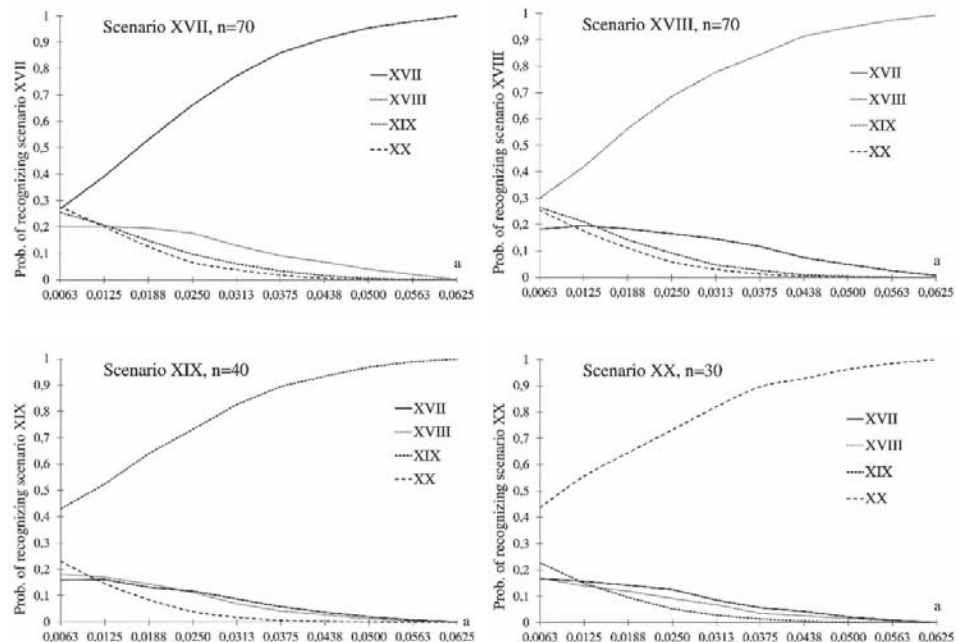
Source: own elaboration.

Figure 9. The PoR actual scenario as one of XIII–XVI scenarios, table 3 × 4



Source: own elaboration.

Figure 10. The PoR actual scenario as one of XVII–XX scenarios, table 4 × 4



Source: own elaboration.

Figures 6–10 show that even when samples are small (e.g. 15 items), probabilities of proper recognition are greater than probabilities of improper recognitions, regardless how small the PFP is. The bigger PFP α , the bigger PoR actual scenario. In the classic statistical testing related to 2×3 CT (see table 6) untrue H_0 has not been rejected even if $n = 60$, PFP $\alpha = 0.05$ and $MoU = 0.2$. In a likelihood based decision dependence between features in 2×3 CT is visible already for $n = 15$ and PFP $\alpha < 0.05$ (see figure 6). In the classic statistical testing related to 2×4 CT (see table 7) untrue H_0 has not been rejected even if $n = 80$, PFP $\alpha = 0.0625$ and $MoU = 0.25$. In a likelihood based decision dependence between features in 2×4 CT is visible already for $n = 25$ and PFP $\alpha = 0.0625$ (see figure 7). In the classic statistical testing related to 3×3 CT (see table 8) untrue H_0 has not been rejected even if $n = 90$, PFP $\alpha = 0.0444$ and $MoU = 0.237$. In a likelihood based decision dependence between features in 3×3 CT is visible already for $n = 30$ and PFP $\alpha < 0.0444$ (see figure 8). A similar situation occurs related to 3×4 and 4×4 CTs.

8.2. Three-way contingency table

Example 4. This example compares decision making by means of classic statistic testing with likelihood based decision making. Tables 11–12 show a set of example three-way CTs filled one by one accordingly to the scenarios XXI–XXVIII. The description of these tables has been presented in the example 1. To decide which of the defined scenarios takes place, see algorithm in section 6.2.

Tables 11–12 show that all the decisions made in a classic way are wrong under the scenarios in question. It is because untrue H_0 hypotheses have not been rejected even in situations where the MoU does not have such small values, e.g. $MoU = 0.267$ in scenario XXVIII. In turn, the parametric approach detected a dependency between features.

Example 5. This example is very similar to the Example 2. The new proposal is a method of statistical inference, not a classical parametric test. All the hypotheses state: “the considered three-way CT is generated accordingly to particular scenarios”. Each of figures from 11 to 12 shows a sets of four likelihood curves. Each curve has its four attributes, namely \hat{a} , n , table dimensions and, of course, the actual generation scenario that is specified in the figure's title. The legend lists all competing scenarios including the actual one.

Table 11. THE CLASSIC STATISTICAL TESTING VERSUS LIKELIHOOD BASED DECISIONS, TABLE 2 x 2 x 2

Experiment settings: $n = 80, \alpha = 0.0625, \alpha = 0.05$

Scenario (MoU)	Content				X and Y dependent	Scenario of maximum likelihood	χ^2	FT	CR	X	H_0 true	H_0 rejected
	5	9	5	14								
XXI(0.130)				14	Yes	XXI	1.505	1.469	1.496	0.979	No	No
	9	15	9	14								
XXII(0.168)	5	5	5	10	Yes	XXII	1.669	1.661	1.665	1.088	No	No
	10	15	15	15								
XXIII(0.188)	5	10	10	10	Yes	XXIII	3.810	4.104	3.855	1.524	No	No
	15	10	10	10								
XXIV(0.250)	5	10	5	10	Yes	XXIV	5.333	5.482	5.350	2.133	No	No
	15	10	15	10								

Source: own elaboration.

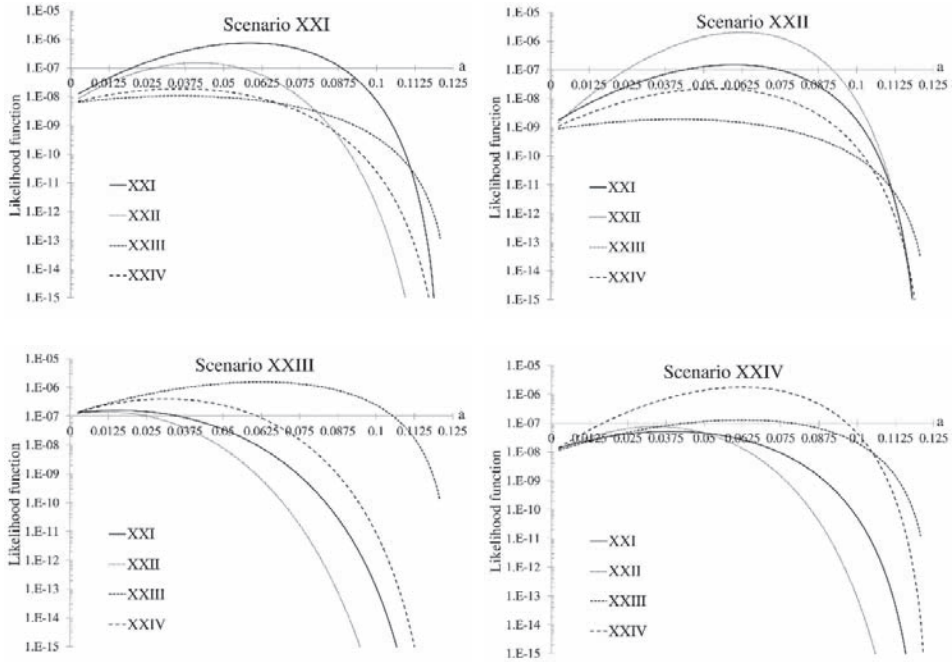
Table 12. THE CLASSIC STATISTICAL TESTING VERSUS LIKELIHOOD BASED DECISIONS, TABLE 3 × 2 × 2

Experiment settings: $n = 120, \alpha = 0.0333, \alpha = 0.05$

Scenario (MoU)	Content			X and Y dependent	Scenario of maximum likelihood	χ^2	FT	CR	X	H_0 true	H_0 rejected
	6	10	6								
XXV(0.089)	6	10	6	Yes	XXV	1.222	1.222	1.221	1.154	No	No
	10	10	10								
	10	14	10								
XXVI(0.149)	6	10	6	Yes	XXVI	3.697	3.629	3.674	1.815	No	No
	10	10	10								
	10	14	10								
XXVII(0.200)	6	14	6	Yes	XXVII	5.000	5.078	5.009	2.500	No	No
	6	14	10								
	10	10	10								
XXVIII(0.267)	6	14	6	Yes	XXVIII	12.800	13.500	12.879	3.200	No	No
	10	10	14								
	14	6	14								

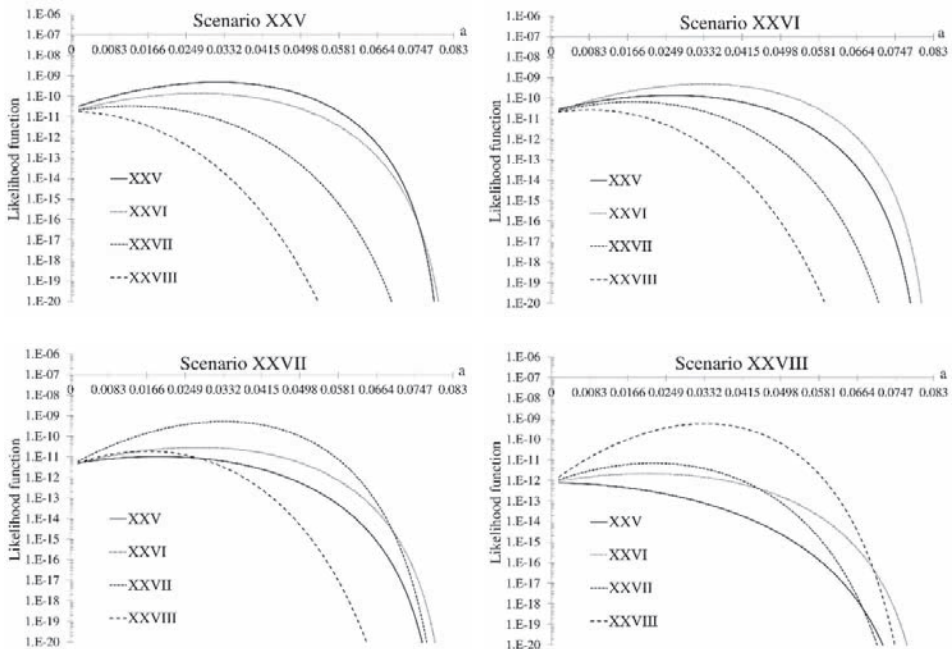
Source: own elaboration.

Figure 11. The likelihood functions for $\hat{a} = 0.0625, n = 80$, table $2 \times 2 \times 2$



Source: own elaboration.

Figure 12. The likelihood functions for $\hat{a} = 0.0333, n = 120$, table $3 \times 2 \times 2$



Source: own elaboration.

It is noteworthy (see figures 11–12) that in each set of likelihood curves, the curve related to the actual scenario predominates over the others. It is instructive to read out a value of a_{max} that maximizes likelihood function on a particular figure and notice that this value is close to assumed \hat{a} .

Example 6. This example is carried out in accordance with the following algorithm:

1. Find dimension of CT in question and read out related q according to the rule:

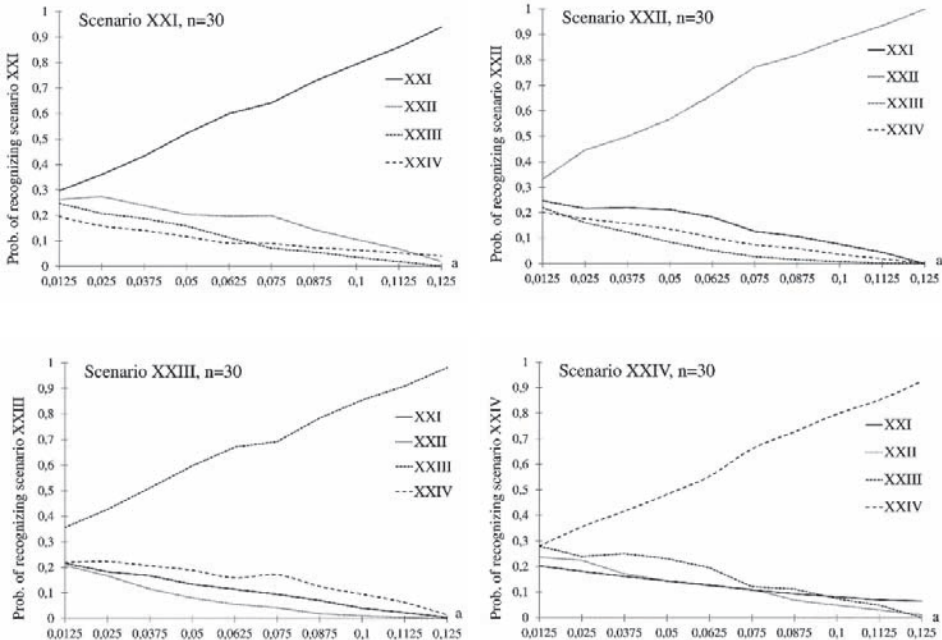
$$2 \times 2 \times 2 (q = 6), 3 \times 2 \times 2 (q = 7)$$

2. Set a set of scenario indices $\{4q - 3, 4q - 2, 4q - 1, 4q\}$.
3. Calculate a value of PFP by means of $a_i = 01i/(wkp)$ ($i = 1, 2, \dots, 10$).

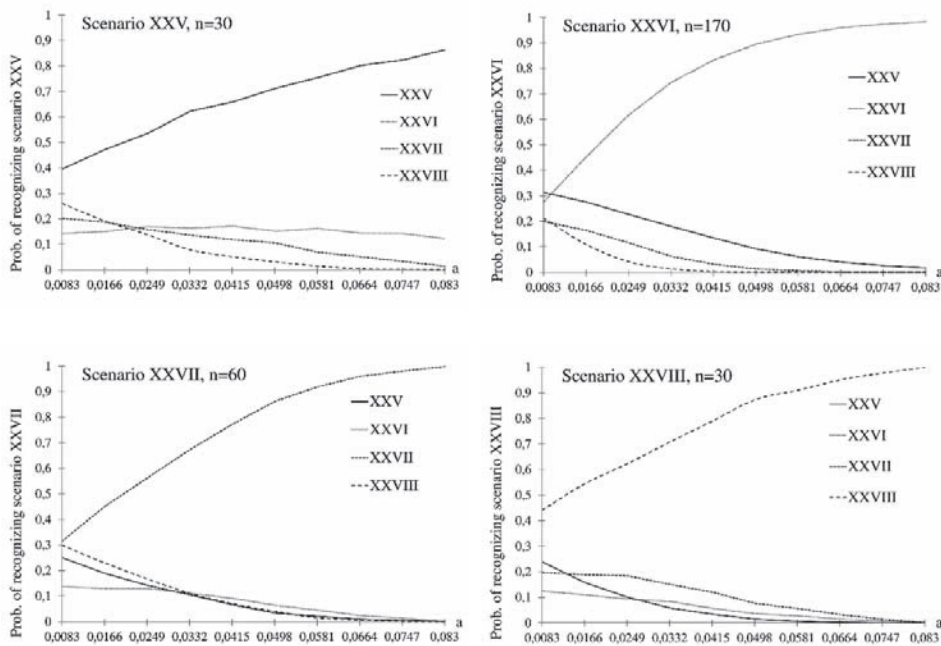
Steps 4–6 are the same as in the example 3.

Figure 13–14 present PoR(a) functions for three-way CT in question. Sample sizes for a given CT are different because a maximal MoU values under the scenarios are different (see table 5). The minimal sample sizes are chosen in such a way that probabilities of proper recognition (actual XXI as XXI, ..., actual XXVIII as XXVIII) are greater than probabilities of improper recognitions for all the a values.

Figure 13. The por actual scenario as one of XXI–XXIV scenarios, table $2 \times 2 \times 2$



Source: own elaboration.

Figure 14. The por actual scenario as one of XXV–XXVIII scenarios, table $3 \times 2 \times 2$ 

Source: own elaboration.

Figures 13–14 show that even when samples are small (e.g. 30 items), probabilities of proper recognition are greater than probabilities of improper recognitions, regardless how small the PFP is. The bigger PFP a , the bigger PoR actual scenario. In the classic statistical testing related to $2 \times 2 \times 2$ CT (see table 11) untrue H_0 has not been rejected even if $n = 80$, PFP $a = 0.0625$ and $MoU = 0.25$. In a likelihood based decision dependence between features in $2 \times 2 \times 2$ CT is visible already for $n = 30$ and PFP $a < 0.0625$ (see figure 13). In the classic statistical testing related to $3 \times 2 \times 2$ CT (see table 12) untrue H_0 has not been rejected even if $n = 120$, PFP $a = 0.0333$ and $MoU = 0.267$. In a likelihood based decision dependence between features in $3 \times 2 \times 2$ CT is visible already for $n = 30$ and PFP $a < 0.0333$ (see figure 14).

9. CONCLUSIONS

There are two new elements in the method of statistical reasoning from CTs presented in this paper. Firstly, CTs are parameterized with the probability flow parameters. Parametric reasoning turns out to be much more sensitive in revealing dependency between features than classic reasoning. Secondly, we suggest a scenario (i.e. internal mechanism) under which particular CT comes into being.

Figuring up more and more general scenarios does not seem very difficult. The researches can be generalized by introducing a part of flow parameter, e.g. $a/2, a/3, \dots$ also remembering about the condition of normalization. The researches can also be generalized by introducing several flow parameters. This, however, causes a significant deterioration in the properties of the parameter estimators. You can always add more parameters to the model, however, this might worsen their estimation. Hence, inflated generalizations should be avoided.

REFERENCES

- Agresti A., (2002), *Categorical Data Analysis*, Wiley, New Jersey.
- Allison P. D., Liker J. K., (1982), *Analyzing Sequential Categorical Data on Dyadic Interaction: A Comment on Gottman*.
- Beh E. J., Davy P. J., (1998), Partitioning Pearson's Chi-squared Statistics for a Completely Ordered Three-Way Contingency Table, *The Australian and New Zealand Journal of Statistics*, 40, 465–477.
- Bishop Y., Fienberg S., Holland, P., (1975), *Discrete Multivariate Analysis – Theory and Practice*, Cambridge, MA: MIT Press.
- Blitzstein J., Diaconis P., (2011), A Sequential Importance Sampling Algorithm for Generating Random Graphs with Prescribed Degrees, *Internet mathematics*, 6 (4), 489–522.
- Bock H. H., (2003), Two-Way Clustering for Contingency Tables: Maximizing A Dependence Measure, *Between data science and applied data analysis*, 143–154.
- Chen Y., Diaconis P., Holmes S. P., Liu J.S., (2005), Sequential Monte Carlo Methods for Statistical Analysis of Tables, *Journal of the American Statistical Association*, 100 (469), 109–120.
- Chen Y., Dinwoodie I. H., Sullivant S., (2006), Sequential Importance Sampling for Multiway Tables, *The Annals of Statistics*, 523–545.
- Cressie N., Read T. R., (1984), Multinomial Goodness-Of-Fit Tests, *Journal of the Royal Statistical Society, Series B (Methodological)*, 440–464.
- Cryan M., Dyer M., (2003), A Polynomial-Time Algorithm to Approximately Count Contingency Tables When the Number of the Rows Is Constant, *Journal of Computer and System Sciences*, 67 (2), 291–310.
- Cryan M., Dyer M., Goldberg L. A., Jerrum M., Martin R., (2006), Rapidly Mixing Markov Chains for Sampling Contingency Tables with A Constant Number of Rows, *SIAM Journal on Computing*, 36 (1), 247–278.
- Cung, B., (2013), *Crime and Demographics: An Analysis of LAPD Crime Data*.
- DeSalvo S., Zhao J. Y., (2015), Random Sampling of Contingency Tables Via Probabilistic Divide-And-Conquer. arXiv preprint arXiv:1507.00070.
- Diaconis P., Sturmfels B., (1998), Algebraic Algorithms for Sampling from Conditional Distributions, *The Annals of statistics*, 26 (1), 363–397.
- Dickhaus T., Straßburger K., Schunk D., Morcillo-Suarez C., Illig T., Navarro A., (2012), How to Analyze Many Contingency Tables Simultaneously in Genetic Association Studies, *Statistical Applications in Genetics and Molecular Biology*, 11 (4), 12.
- El Galta R., Stijnen T., Houwing-Duistermaat J. J., (2008), Testing for Genetic Association: A Powerful Score Test, *Statistics in Medicine*, 27 (22), 4596–4609.
- Fishman G. S., (2012), Counting Contingency Tables Via Multistage Markov Chain Monte Carlo. *Journal of Computational and Graphical Statistics*, 21(3), 713–738.
- Freeman M. F., Tukey J. W., (1950), Transformations Related to the Angular and the Square Root, *Annals of Mathematical Statistics*, 21, 607–611.

- Gokhale D., Kullback S., (1978), *The Information in Contingency Tables*, Marcel Dekker, Inc., New York.
- Goodman L., Kruskal W., (1954), Measures of Association for Cross Classifications, *Journal of the American Statistical Association*, 49, 732–764.
- Gray L. N., Williams J. S., (1975), Goodman and Kruskal's Tau b: Multiple and Partial Analogs, In: Proceedings of the Social Statistics Section, *Journal of the American Statistical Association*, 444–448.
- Haas P. J., Hueske F., Markl V., (2007), Detecting Attribute Dependencies from Query Feedback, Proceedings of the 33rd International Conference on Very Large Data Bases, VLDB Endowment, 830–841.
- Harshman R. A., (1970), Foundations of the PARAFAC Procedure: Models and Conditions for an Explanatory Multi-modal Factor Analysis, UCLA Working Papers in Phonetics, 16, 1–84.
- Ilyas I. F., Markl V., Haas P., Brown P., Aboulnaga A., (2004), CORDS: Automatic Discovery of Correlations and Soft Functional Dependencies, Proceedings of the 2004 ACM SIGMOD international conference on Management of data, ACM, 647–658.
- Iossifova R., Marmolejo-Ramos F., (2013), When The Body Is Time: Spatial and Temporal Deixis in Children with Visual Impairments and Sighted Children, *Research in Developmental Disabilities*, 34 (7), 2173–2184.
- Kaski S., Nikkila J., Sinkkonen J., Lahti L., Knuutila J. E., Roos C., (2005), Associative Clustering for Exploring Dependencies Between Functional Genomics Data Sets, *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 2 (3), 203–216.
- Kullback S., (1959), *Information Theory and Statistics*, Wiley, New York.
- Lombardo R., (2011), Three-Way Association Measure Decompositions: The Delta index, *Journal of Statistical Planning and Inference*, 141, 1789–1799.
- Lombardo R., Beh E. J., (2010), Simple and Multiple Correspondence Analysis for Ordinal-Scale Variables Using Orthogonal Polynomials, *Journal of Applied Statistics*, 37 (12), 2101–2116.
- Marcotorchino F., (1984), *Utilisation des Comparaisons par Paires en Statistique des Contingences. Partie (I)*, Etude IBM F069, France.
- Neyman J., (1949), Contribution to the Theory of the χ^2 Test, Proceedings of the First Berkeley Symposium on Mathematical Statistics and Probability, Univ. of Calif. Press, 239–273.
- Oates T., Cohen P. R., (1996), Searching for Structure in Multiple Streams of Data, ICML 96.
- Pardo M. C., (1996), An Empirical Investigation of Cressie and Read Tests for the Hypothesis of Independence in Three-way Contingency Tables, *Kybernetika* 32 (2), 175–183.
- Pearson K., (1900), On the Criterion That a Given System of Deviations From the Probable in the Case of a Correlated System of Variables is Such That it Can be Reasonably Supposed to Have Arisen From Random Sampling, *Philosophy Magazine Series*, 50, 157–172.
- Schrepp M., (2003), A Method for The Analysis of Hierarchical Dependencies Between Items of a Questionnaire, *Methods of Psychological Research Online*, 19, 43–79.
- Sokal R. R., Rohlf F. J., (2012), *Biometry: The Principles and Practice of Statistics in Biological Research*, Freeman, New York.
- Steinle M., Aberer K., Girdzijauskas S., Lovis C., (2006), Mapping Moving Landscapes by Mining Mountains of Logs: Novel Techniques for Dependency Model Generation, Proceedings of the 32nd international conference on Very large data bases, VLDB Endowment, 1093–1102.
- Sulewski P., (2009), Two-By-Two Contingency Table as a Goodness-Of-Fit Test, *Computational Methods in Science and Technology*, 15 (2), 203–211.
- Sulewski P., (2013), Modyfikacja testu niezależności, *Wiadomości Statystyczne*, 10, 1–19.
- Sulewski P., (2016), Moc testów niezależności w tablicy dwudzielczej większej niż 2x2, *Przegląd Statystyczny*, 63 (2), 191–209.
- Sulewski P., (2018a), Power Analysis of Independence Testing for the Three-way Contingency Tables of Small Sizes, *Journal of Applied Statistics*, 45 (13), 2481–2498.
- Sulewski P., (2018b), Nonparametric Versus Parametric Reasoning Based on 2x2 Contingency Tables, *Computational Methods in Science and Technology*, 24 (2), 143–153.

- Sulewski P., Motyka R., (2015), Power Analysis of Independence Testing for Contingency Tables. *Scientific Journal of Polish Naval Academy*, 56, 37–46.
- Tucker L. R., (1963), Implications of Factor Analysis of Three-Way Matrices for Measurement of Change. *In Problems in Measuring Change*, 122–137.
- Van Belle G., Fisher L.D., Heagerty P. J., Lumley T., (2004), *Categorical Data: Contingency Tables*, John Wiley & Sons.
- Yoshida R., Xi J., Wei S., Zhou F., Haws D., (2011), Semigroups and Sequential Importance Sampling for Multiway Tables. arXiv preprint arXiv:1111.6518.

WNIOSKOWANIE PARAMETRYCZNE I NIEPARAMETRYCZNE W TABLICACH DWUDZIELCZYCH I TRÓJDZIELCZYCH

Streszczenie

W artykule proponowane są scenariusze generowania tablic dwudzielczych (TD) z parametrem przepływu prawdopodobieństwa i zdefiniowane są miary nieprawdliwości H_0 . W artykule wykorzystywane są statystyki z rodziny X^2 oraz statystyka modułowa $|X|$. Niniejsza praca jest prostą próbą zastąpienia nieparametrycznej metody wnioskowania statystycznego metodą parametryczną. Metoda największej wiarygodności jest wykorzystana do oszacowania parametru przepływu prawdopodobieństwa. W pracy opisane są także instrukcje generowania TD za pomocą metody słupkowej. Symulacje komputerowe przeprowadzono metodami Monte Carlo.

Słowa kluczowe: wnioskowanie statystyczne, funkcja największej wiarygodności, tablice kontyngencji, test parametryczny, parametr przepływu prawdopodobieństwa

NONPARAMETRIC VERSUS PARAMETRIC REASONING BASED ON TWO-WAY AND THREE-WAY CONTINGENCY TABLES

Abstract

This paper proposes scenarios of generating two-way and three way contingency tables (CTs). A concept of probability flow parameter (PFP) plays a crucial role in these scenarios. Additionally, measures of untruthfulness of H_0 are defined. The power divergence statistics and the $|X|$ statistics are used. This paper is a simple attempt to replace a nonparametric statistical inference from CTs by the parametric one. Maximum likelihood method is applied to estimate PFP and instructions of generating CTs according to scenarios in question are presented. The Monte Carlo method is used to carry out computer simulations.

Keywords: statistical inference, likelihood function, contingency tables, parametric test, probability flow parameter