Noname manuscript No.
(will be inserted by the editor)

Familywise decompositions of Pearson's chi-square statistic in the analysis of contingency tables

Rosaria Lombardo · Yoshio Takane · Eric J Beh

the date of receipt and acceptance should be inserted later

Abstract Pearson's chi-square statistic is well established for testing goodness-of-fit of various hypotheses about observed frequency distributions in contingency tables. A general formula for ANOVA-like decompositions of Pearson's statistic is given under the independence assumption along with their extensions to higher-order tables. Mathematically, it makes the terms in the partitions and orthogonality among them obvious. Practically, it enables simultaneous analyses of marginal and joint probabilities in contingency tables under a variety of hypotheses about the marginal probabilities. Specifically, this framework accommodates the specification of theoretically driven probabilities as well as the well known cases in which the marginal probabilities are fixed or estimated from the data. The former allows tests of prescribed marginal probabilities, while the latter allows tests of the associations among variables after eliminating the marginal effects. Mixtures of these two cases are also permitted. Examples are given to illustrate the tests.

Subject classification JEL C25, AMS 15A03 15A09

Keywords Pearson's chi-square statistic · Orthogonal projectors · Familywise decompositions · Hypothesised probabilities · Observed proportions

R. Lombardo
Economics Department, University of Campania "Luigi Vanvitelli", Italy
Tel.: +39-0823-274032
Fax: +39-0823-622984
E-mail: rosaria.lombardo@unicampania.it

Y. Takane
Department of Psychology, University of Victoria, Canada
E-mail: yoshio.takane@mcgill.ca

E. J. Beh
School of Mathematical & Physical Sciences, University of Newcastle, Australia
E-mail: eric.beh@newcastle.edu.au

1 Introduction

The analysis of multiple categorical variables is a cornerstone in many scientific investigations. Measuring and assessing the nature of the association among these variables are often undertaken by partitioning a multi-way statistic into univariate, bivariate, trivariate, and higher order terms (Carlier and Kroonenberg, 1996; Lombardo et al., 1996; Loisel and Takane, 2016; Beh and Davy, 1998). For three-way contingency tables, Lancaster (1951) presented two methods of partitioning Pearson's chi-square statistic (Pearson, 1900). One, described in Section 4 of his paper, is based on repeated calculations of the chi-square statistics for full and various marginal tables and subtracting the latter from the former. An advantage of his partition is that it provides flexibility in specifying *a priori* marginal probabilities. His exposition was rather informal, however, and no explicit expressions of the terms in the partition were given, nor proofs of their orthogonality; see also Bishop et al. (1975, p.361). Toward the end of his paper, Lancaster proposed another method, using orthogonal transformations of variables arranged in hierarchical order. This method also enjoys flexibility in specifying *a priori* marginal probabilities. His discussion, however, focused primarily on the analysis of a $2 \times 2 \times 2$ table. Although he claimed that it could be implemented for the analysis of variables consisting of more than 2 categories, he did not present any formal derivations, proofs, or examples, to highlight how this could be achieved.

In this paper, we give a new general framework and strategy for orthogonal partitions of Pearson's chi-square statistic under the assumption of complete independence of the variables. This method, along with matrix formulations, overcomes some of the difficulties of Lancaster's methods mentioned above. Specifically, it provides explicit expressions of the terms in the partitions, which makes their meaning (i.e., what they actually represent) clearer. This method makes orthogonality and other relations (e.g., inclusion) among the terms in the partitions easier to observe. It is also applicable to contingency tables of any size, and is easily extensible to tables of any order.

Another advantage of the proposed partition is that it accommodates a variety of different hypotheses concerning the distribution of the marginal probabilities (expected proportions). Simultaneous tests of the marginal and joint probabilities are possible (Lang, 1996). The partitions presented in this paper are based on the ANOVA-like familywise partitions of Pearson's chi-square statistic (Pearson, 1900; Lancaster, 1951) when alternative representations of the expected proportions (hypothesised probabilities) are provided. Most often, these probabilities are estimated by using the margins of the empirical distribution that underly the data (Carlier and Kroonenberg, 1996; Kroonenberg, 2008; Lombardo et al., 1996; Beh et al., 2007; Beh and Lombardo, 2014). In some cases, however, the probabilities are prescribed by the analyst, because *a priori* knowledge of the phenomena suggests otherwise (Loisel and Takane, 2016). See the discussion of Andersen (1980, p.92-93) when, for example, it is known that all categories have *a priori* the same probability of occurring. Also, much like goodness-of-fit tests designed for the analysis of one-way data where the hypothesised marginal probabilities are theoretically derived (Agresti, 1990; Andersen, 1980, 1991), studying two-way and multi-way tables in a similar manner poses interesting results (as we shall see).

A variety of other methods and approaches have been proposed for analyses of contingency tables (Agresti, 1990; Andersen, 1980, 1991; Bishop et al., 1975; Friendly, 1994; Goodman, 1969, 1970; van der Heijden and de Leeuw, 1985). Among them, perhaps the strongest competitor to Pearson's chi-square statistic is the log-likelihood ratio (LR) statistic often used in log-linear analyses of contingency tables. Partitions similar to those presented in the present paper for Pearson's statistic have also been derived for the LR statistic (Cheng et al., 2006; Goodman, 1969; Loisel and Takane, 2016). We discuss these partitions and compare them with those for Pearson's statistic in Section 4.1, using a concrete example.

This paper is organized as follows. In Section 2, we present formal derivations of the basic partitions. We first give some general results on orthogonal projectors to be used in the derivations (Section 2.1). We then apply these results to derive partitions for one-way (Section 2.2), two-way (Section 2.3), three-way (Section 2.4), and higher-order tables (Section 2.5). In deriving these partitions, we also discuss their basic properties (e.g., orthogonality of terms in the partitions). In Section 3, we discuss representative scenarios concerning possible specifications of marginal probabilities, and their consequences in distributional properties of the terms in the partitions. Section 4 provides some empirical examples to demonstrate the use of the proposed partitions, followed by brief concluding remarks in Section 5.

2 Partitioning the chi-square statistic

In this section, we give ANOVA-like familywise partitions of Pearson's chi-square statistic under arbitrary marginal probabilities. We start with one-way tables, and then extend the results obtained for one-way tables to higher-order tables. This makes it easier to see that our approach can be easily extended to any complex contingency tables.

2.1 Preliminary results

Consider a categorical variable, A, with A categories belonging to the space $\mathcal{R}^A$, and let $\mathbf{f}_A$ denote an A-component vector of observed frequencies from a multinomial population $\mathrm{MN}(\mathbf{p}_A)$ out of $n = \mathbf{f}'_A\mathbf{1}_A$ independently replicated trials, where $\mathbf{p}_A$ is the vector of probabilities of A mutually exclusive events, and $\mathbf{1}_A$ is the A-component vector of unities. Let $\mathbf{D}_A$ denote the diagonal matrix with the elements of $\mathbf{p}_A$ as its diagonal elements so that $\mathbf{p}_A = \mathbf{D}_A\mathbf{1}_A$, and let $\hat{\mathbf{p}}_A = \mathbf{f}_A/n$ be an observed counterpart to $\mathbf{p}_A$. Pearson's chi-square statistic in this situation can be expressed as

$$
\begin{aligned}
X^2 &= n(\hat{\mathbf{p}}_A - \mathbf{p}_A)'\mathbf{D}_A^{-1}(\hat{\mathbf{p}}_A - \mathbf{p}_A) \\
&= n(\hat{\mathbf{p}}_A - \mathbf{p}_A)'\mathbf{D}_A^{-1}\hat{\mathbf{p}}_A = n(\hat{\mathbf{p}}_A'\mathbf{D}_A^{-1}\hat{\mathbf{p}}_A - 1),
\end{aligned} \tag{1}
$$

which is known to asymptotically follow the chi-square distribution with $A-1$ degrees of freedom (df) (written as $X^2 \rightsquigarrow \chi^2_{A-1}$) when the specified value of $\mathbf{p}_A = \mathbf{p}_A^*$ is correct. Note that $\mathbf{p}_A$ is the expected value of $\hat{\mathbf{p}}_A$, and that $\mathbf{D}_A^{-1}$ is a g-inverse of the covariance matrix of $\sqrt{n}\hat{\mathbf{p}}_A$, given by $n\mathrm{Var}[\hat{\mathbf{p}}_A] = \mathbf{D}_A - \mathbf{p}_A\mathbf{p}_A' \equiv \boldsymbol{\Sigma}_A$. It is interesting to note that $\mathbf{D}_A^{-1}$ in (1) can be replaced by any g-inverse matrix of $\boldsymbol{\Sigma}_A$, with $\mathbf{D}_A^{-1}$ being a special case.

The following projectors play central roles in the sequel:

$$
\mathbf{R}_{1/A} = \mathbf{D}_A^{1/2}\mathbf{1}_A\mathbf{1}_A'\mathbf{D}_A^{1/2} = \mathbf{p}_A^{1/2}(\mathbf{p}_A^{1/2})', \tag{2}
$$

and its complement,

$$
\mathbf{Q}_{1/A} = \mathbf{I}_A - \mathbf{R}_{1/A} = \mathbf{I}_A - \mathbf{D}_A^{1/2}\mathbf{1}_A\mathbf{1}_A'\mathbf{D}_A^{1/2} = \mathbf{I}_A - \mathbf{p}_A^{1/2}(\mathbf{p}_A^{1/2})', \tag{3}
$$

where $\mathbf{I}_A$ is the identity matrix of order A. Note that $\mathbf{R}_{1/A}$ is the orthogonal projector onto $\mathrm{Sp}(\mathbf{D}_A^{1/2}\mathbf{1}_A)$, while $\mathbf{Q}_{1/A}$ is the orthogonal projector onto $\mathrm{Ker}(\mathbf{1}_A'\mathbf{D}_A^{1/2})$, where Sp indicates the space spanned by the column vectors in its argument, and Ker indicates the null space of its argument, i.e., $\mathrm{Ker}(\mathbf{1}_A'\mathbf{D}_A^{1/2})$ indicates the space spanned by all vectors $\mathbf{x}$'s such that $\mathbf{1}_A'\mathbf{D}_A^{1/2}\mathbf{x} = 0$. It follows that

$$
\mathbf{R}_{1/A} + \mathbf{Q}_{1/A} = \mathbf{I}_A \quad \text{(complementarity)}, \tag{4}
$$

$$
\mathbf{R}_{1/A}^2 = \mathbf{R}_{1/A} \text{ and } \mathbf{Q}_{1/A}^2 = \mathbf{Q}_{1/A} \quad \text{(idempotency)}, \tag{5}
$$

$$
\mathbf{R}_{1/A}' = \mathbf{R}_{1/A} \text{ and } \mathbf{Q}_{1/A}' = \mathbf{Q}_{1/A} \quad \text{(symmetry)}, \tag{6}
$$

and

$$
\mathbf{R}_{1/A}\mathbf{Q}_{1/A} = \mathbf{Q}_{1/A}\mathbf{R}_{1/A} = \mathbf{O}_A, \quad \text{(orthogonality)}, \tag{7}
$$

where $\mathbf{O}_A$ is the zero matrix of order A. These properties are standard properties of orthogonal projectors (Yanai et al., 2011), and can easily be verified directly. Note that $\boldsymbol{\Sigma}_A$ introduced earlier can be expressed as $\boldsymbol{\Sigma}_A = \mathbf{D}_A^{1/2}\mathbf{Q}_{1/A}\mathbf{D}_A^{1/2}$.

2.2 One-way tables

When applying the complementarity property of Equation (4) to $\mathbf{D}_A^{-1/2}\hat{\mathbf{p}}_A$, we obtain

$$
\mathbf{D}_A^{-1/2}\hat{\mathbf{p}}_A = (\mathbf{R}_{1/A} + \mathbf{Q}_{1/A})\mathbf{D}_A^{-1/2}\hat{\mathbf{p}}_A = \mathbf{R}_{1/A}\mathbf{D}_A^{-1/2}\hat{\mathbf{p}}_A + \mathbf{Q}_{1/A}\mathbf{D}_A^{-1/2}\hat{\mathbf{p}}_A. \tag{8}
$$

The orthogonality of the two terms on the right hand side of (8) is obvious from Equation (7). The n times squared norm of the second term is given by

$$\begin{aligned} X^2_{Total} \equiv X^2_A &= n\hat{\mathbf{p}}'_A\mathbf{D}_A^{-1/2}\mathbf{Q}_{1/A}\mathbf{D}_A^{-1/2}\hat{\mathbf{p}}_A \\ &= n\hat{\mathbf{p}}'_A(\mathbf{D}_A^{-1} - \mathbf{1}_A\mathbf{1}'_A)\hat{\mathbf{p}}_A = n(\hat{\mathbf{p}}'_A\mathbf{D}_A^{-1}\hat{\mathbf{p}}_A - 1), \end{aligned} \tag{9}$$

which is equal to X^2 in Equation (1). It is interesting to note that $\mathbf{D}_A^{-1} - \mathbf{1}_A\mathbf{1}'_A$ in the above formula is also a (symmetric reflexive) g-inverse of $\boldsymbol{\Sigma}_A$.

Let p_a and $\hat{p}_a$ $(a = 1, \cdots, A)$ denote the a^{th} element of $\mathbf{p}_A$ and $\hat{\mathbf{p}}_A$, respectively. Then Equation (9) can be rewritten, in scalar notation, as

$$X^2_{Total} = n\sum_{a=1}^{A}\frac{1}{p_a}(\hat{p}_a - p_a)^2 = n\left(\sum_{a=1}^{A}\frac{\hat{p}_a^2}{p_a} - 1\right). \tag{10}$$

2.3 Two-way tables

Suppose that there is a second categorical variable, B, with B categories. Let $\hat{\mathbf{P}}_{AB}$ denote an A by B table of observed probabilities of cross classified events by the two variables. Let $\mathbf{D}_A$ and $\mathbf{D}_B$ denote the diagonal matrices whose diagonal elements are, respectively, the elements of $\mathbf{p}_A$ and $\mathbf{p}_B$, the vectors of marginal probabilities of variables A and B. Furthermore, let $\mathbf{R}_{1/B}$ and $\mathbf{Q}_{1/B}$ be defined analogously to $\mathbf{R}_{1/A}$ and $\mathbf{Q}_{1/A}$, as in Equations (2) and (3). These matrices have similar properties to $\mathbf{R}_{1/A}$ and $\mathbf{Q}_{1/A}$, as stated in Equation (4) through Equation (7).

For the purpose of partitioning onto the space $\mathcal{R}^{A\times B}$, we apply $\mathbf{R}_{1/A}+\mathbf{Q}_{1/A}$ to the left hand side and $\mathbf{R}_{1/B}+\mathbf{Q}_{1/B}$ to the right hand side of $\mathbf{D}_A^{-1/2}\hat{\mathbf{P}}_{AB}\mathbf{D}_B^{-1/2}$ to obtain

$$\mathbf{D}_A^{-1/2}\hat{\mathbf{P}}_{AB}\mathbf{D}_B^{-1/2} = (\mathbf{R}_{1/A} + \mathbf{Q}_{1/A})\mathbf{D}_A^{-1/2}\hat{\mathbf{P}}_{AB}\mathbf{D}_B^{-1/2}(\mathbf{R}_{1/B} + \mathbf{Q}_{1/B}), \tag{11}$$

which can be vectorized, using a vec operator and Kronecker products, as

$$\begin{aligned} \mathbf{M}_{AB}^{1/2}\hat{\mathbf{p}}_{AB} &\equiv \mathrm{vec}(\mathbf{D}_A^{-1/2}\hat{\mathbf{P}}_{AB}\mathbf{D}_B^{-1/2}) \\ &= [(\mathbf{R}_{1/B} + \mathbf{Q}_{1/B}) \otimes (\mathbf{R}_{1/A} + \mathbf{Q}_{1/A})]\mathbf{M}_{AB}^{1/2}\hat{\mathbf{p}}_{AB} \\ &= (\mathbf{R}_{1/B} \otimes \mathbf{R}_{1/A} + \mathbf{R}_{1/B} \otimes \mathbf{Q}_{1/A} + \mathbf{Q}_{1/B} \otimes \mathbf{R}_{1/A} \\ &\qquad + \mathbf{Q}_{1/B} \otimes \mathbf{Q}_{1/A})\mathbf{M}_{AB}^{1/2}\hat{\mathbf{p}}_{AB}, \end{aligned} \tag{12}$$

where $\mathbf{M}_{AB} = \mathbf{D}_B^{-1} \otimes \mathbf{D}_A^{-1}$. (Incidentally, this matrix is a g-inverse of the covariance matrix of $\sqrt{n}\hat{\mathbf{p}}_{AB}$, which is given by $\boldsymbol{\Sigma}_{AB} = \mathbf{D}_B \otimes \mathbf{D}_A - \mathbf{p}_B\mathbf{p}'_B \otimes \mathbf{p}_A\mathbf{p}'_A$.)

The total chi-square is calculated by

$$\begin{aligned} X^2_{Total} &= n(\hat{\mathbf{p}}_{AB} - \mathbf{p}_B \otimes \mathbf{p}_A)'\mathbf{M}_{AB}(\hat{\mathbf{p}}_{AB} - \mathbf{p}_B \otimes \mathbf{p}_A) \\ &= n\hat{\mathbf{p}}'_{AB}\mathbf{M}_{AB}^{1/2}[\mathbf{I}_{AB} - (\mathbf{R}_{1/B} \otimes \mathbf{R}_{1/A})]\mathbf{M}_{AB}^{1/2}\hat{\mathbf{p}}_{AB} \\ &= n\hat{\mathbf{p}}'_{AB}\mathbf{M}_{AB}^{1/2}\mathbf{Q}_{1/AB}\mathbf{M}_{AB}^{1/2}\hat{\mathbf{p}}_{AB}, \end{aligned} \tag{13}$$

where $\mathbf{Q}_{1/AB} = \mathbf{I}_{AB} - \mathbf{R}_{1/AB}$ and $\mathbf{R}_{1/AB} = \mathbf{R}_{1/B} \otimes \mathbf{R}_{1/A}$. Notice that $\mathbf{R}_{1/AB}$ and $\mathbf{Q}_{1/AB}$ are both orthogonal projectors having similar properties to $\mathbf{R}_{1/A}$ and $\mathbf{Q}_{1/B}$, as given in Equations (4) through (7).

The orthogonality of the four terms on the right hand side of Equation (12) may be readily seen from Equation (7). For example, the second and the fourth terms are orthogonal, since

$$\begin{aligned}
&\hat{\mathbf{p}}'_{AB}\mathbf{M}_{AB}^{1/2}(\mathbf{R}_{1/B} \otimes \mathbf{Q}_{1/A})(\mathbf{Q}_{1/B} \otimes \mathbf{Q}_{1/A})\mathbf{M}_{ab}^{1/2}\hat{\mathbf{p}}_{AB} \\
&= \hat{\mathbf{p}}'_{AB}\mathbf{M}_{AB}^{1/2}(\mathbf{R}_{1/B}\mathbf{Q}_{1/B} \otimes \mathbf{Q}_{1/A})\mathbf{M}_{AB}^{1/2}\hat{\mathbf{p}}_{AB} \\
&= \hat{\mathbf{p}}'_{AB}\mathbf{M}_{AB}^{1/2}(\mathbf{O}_{B} \otimes \mathbf{Q}_{1/A})\mathbf{M}_{AB}^{1/2}\hat{\mathbf{p}}_{AB} = 0,
\end{aligned} \tag{14}$$

where $\mathbf{O}_B$ is the zero matrix of order B. Consequently, X^2_{Total} can be partitioned as

$$X^2_{Total} = X^2_A + X^2_B + X^2_{AB}, \tag{15}$$

where

$$X^2_A = n\hat{\mathbf{p}}'_{AB}\mathbf{M}_{AB}^{1/2}(\mathbf{R}_{1/B} \otimes \mathbf{Q}_{1/A})\mathbf{M}_{AB}^{1/2}\hat{\mathbf{p}}_{AB}, \tag{16}$$

$$X^2_B = n\hat{\mathbf{p}}'_{AB}\mathbf{M}_{AB}^{1/2}(\mathbf{Q}_{1/B} \otimes \mathbf{R}_{1/A})\mathbf{M}_{AB}^{1/2}\hat{\mathbf{p}}_{AB}, \tag{17}$$

and

$$X^2_{AB} = n\hat{\mathbf{p}}'_{AB}\mathbf{M}_{AB}^{1/2}(\mathbf{Q}_{1/B} \otimes \mathbf{Q}_{1/A})\mathbf{M}_{AB}^{1/2}\hat{\mathbf{p}}_{AB}. \tag{18}$$

The X^2_A, X^2_B, and X^2_{AB} represent part chi-squares for the main effect of A, the main effect of B, and the interaction between A and B.

Let $\mathbf{Q}_X$ and $\mathbf{Q}_Y$ be two orthogonal projectors such that $\mathbf{Q}_X\mathbf{Q}_Y = \mathbf{Q}_Y\mathbf{Q}_X = \mathbf{Q}_Y$. This implies $\mathrm{Sp}(\mathbf{Q}_X) \supset \mathrm{Sp}(\mathbf{Q}_Y)$, that is, the latter is a subspace of the former. We have $\mathbf{Q}_{1/AB}(\mathbf{R}_{1/B} \otimes \mathbf{Q}_{1/A}) = (\mathbf{R}_{1/B} \otimes \mathbf{Q}_{1/A})\mathbf{Q}_{1/AB} = \mathbf{R}_{1/B} \otimes \mathbf{Q}_{1/A}$, $\mathbf{Q}_{1/AB}(\mathbf{Q}_{1/B} \otimes \mathbf{R}_{1/A}) = (\mathbf{Q}_{1/B} \otimes \mathbf{R}_{1/A})\mathbf{Q}_{1/AB} = \mathbf{Q}_{1/B} \otimes \mathbf{R}_{1/A}$, and $\mathbf{Q}_{1/AB}(\mathbf{Q}_{1/B} \otimes \mathbf{Q}_{1/A}) = (\mathbf{Q}_{1/B} \otimes \mathbf{Q}_{1/A})\mathbf{Q}_{1/AB} = \mathbf{Q}_{1/B} \otimes \mathbf{Q}_{1/A}$, so that $\mathrm{Sp}(\mathbf{Q}_{1/AB}) \supset \mathrm{Sp}(\mathbf{R}_{1/B} \otimes \mathbf{Q}_{1/A})$, $\mathrm{Sp}(\mathbf{Q}_{1/AB}) \supset \mathrm{Sp}(\mathbf{Q}_{1/B} \otimes \mathbf{R}_{1/A})$, and $\mathrm{Sp}(\mathbf{Q}_{1/AB}) \supset \mathrm{Sp}(\mathbf{Q}_{1/B} \otimes \mathbf{Q}_{1/A})$, implying that $\mathrm{Sp}(\mathbf{Q}_{1/AB})$ includes all of the other three spaces as its subspaces.

Let $p_{a.}$ and $p_{.b}$ be the a^{th} and b^{th} elements of $\mathbf{p}_A$ and $\mathbf{p}_B$, and let $\hat{p}_{a.}$ and $\hat{p}_{.b}$ represent their observed counterparts. Further, let $\hat{p}_{ab}$ represent the observed joint proportion of the a^{th} category of variable A and the b^{th} category of variable B (i.e., the ab^{th} element of $\hat{\mathbf{P}}_{AB}$). Then, X^2_{Total}, X^2_A, X^2_B, and X^2_{AB} can be written as

$$X^2_{Total} = n\sum_{a=1}^{A}\sum_{b=1}^{B}\frac{1}{p_{a.}p_{.b}}(\hat{p}_{ab} - p_{a.}p_{.b})^2, \tag{19}$$

$$X^2_A = n\sum_{a=1}^{A}\frac{1}{p_{a.}}(\hat{p}_{a.} - p_{a.})^2, \tag{20}$$

$$X^2_B = n\sum_{b=1}^{B}\frac{1}{p_{.b}}(\hat{p}_{.b} - p_{.b})^2, \tag{21}$$

and

$$X^2_{AB} = n \sum_{a=1}^{A} \sum_{b=1}^{B} \frac{1}{p_{a.}p_{.b}} (\hat{p}_{ab} - \hat{p}_{a.}p_{.b} - p_{a.}\hat{p}_{.b} + p_{a.}p_{.b})^2. \tag{22}$$

The correspondence between the matrix and scalar representations can be easily verified by elaborating the matrices between $\hat{\mathbf{p}}'_{AB}$ and $\hat{\mathbf{p}}_{AB}$ in Equations (13), (16), (17), and (18). Note that Equations (18) and (22) provide explicit expressions of X^2_{AB}, which were defined by

$$X^2_{AB} = X^2_{Total} - X^2_A - X^2_B$$

in Lancaster (1951) based on the assumption of Equation (15).

2.4 Three-way tables

Suppose that, in addition to the previous two variables, there is a third variable, called C with C categories, giving rise to a three-way contingency tables. Let $\hat{\mathbf{p}}_{ABC}$ denote a vector of observed proportions $\hat{p}_{abc}$ of the joint event abc arranged in such a way that a is the fastest moving index and c is the slowest moving index. Let $\mathbf{D}_A$, $\mathbf{D}_B$, and $\mathbf{D}_C$ represent the diagonal matrices whose diagonal elements are marginal probabilities of categories of the three variables, and let $\mathbf{M}_{ABC} = \mathbf{D}_C^{-1} \otimes \mathbf{D}_B^{-1} \otimes \mathbf{D}_A^{-1}$. Define $\mathbf{R}_{1/C}$ and $\mathbf{Q}_{1/C}$ analogously to $\mathbf{R}_{1/A}$ and $\mathbf{Q}_{1/A}$ as in Equations (2) and (3). Then, the projection of $\mathbf{M}^{1/2}_{ABC}\hat{\mathbf{p}}_{ABC}$ onto the space $\mathcal{R}^{A\times B\times C}$ is given by

$$\begin{aligned}
&\mathbf{M}^{1/2}_{ABC}\hat{\mathbf{p}}_{ABC}\\
&\quad= [(\mathbf{R}_{1/C} + \mathbf{Q}_{1/C}) \otimes (\mathbf{R}_{1/B} + \mathbf{Q}_{1/B}) \otimes (\mathbf{R}'_{1/A} + \mathbf{Q}_{1/A})]\mathbf{M}^{1/2}_{ABC}\hat{\mathbf{p}}_{ABC}\\
&\quad= (\mathbf{R}_{1/C} \otimes \mathbf{R}_{1/B} \otimes \mathbf{R}_{1/A} + \mathbf{R}_{1/C} \otimes \mathbf{R}_{1/B} \otimes \mathbf{Q}_{1/A} + \mathbf{R}_{1/C} \otimes \mathbf{Q}_{1/B} \otimes \mathbf{R}_{1/A}\\
&\quad+\mathbf{Q}_{1/C} \otimes \mathbf{R}_{1/B} \otimes \mathbf{R}_{1/A} + \mathbf{R}_{1/C} \otimes \mathbf{Q}_{1/B} \otimes \mathbf{Q}_{1/A} + \mathbf{Q}_{1/C} \otimes \mathbf{R}_{1/B} \otimes \mathbf{Q}_{1/A}\\
&\qquad\quad+\mathbf{Q}_{1/C} \otimes \mathbf{Q}_{1/B} \otimes \mathbf{R}_{1/A} + \mathbf{Q}_{1/C} \otimes \mathbf{Q}_{1/B} \otimes \mathbf{Q}_{1/A})\mathbf{M}^{1/2}_{ABC}\hat{\mathbf{p}}_{ABC}.
\end{aligned} \tag{23}$$

The eight terms on the righthand side of Equation (23) are mutually orthogonal, which can be readily verified in a manner similar to Equation (14). The n times squared norms of the last seven terms give part chi-squares for the three main effects A, B, and C, the three two-way interaction effects AB, AC, and BC, and the one three-way interaction effect ABC. Specifically, we can write the n times squared norm of $\mathbf{M}^{1/2}_{ABC}\hat{\mathbf{p}}_{ABC}$ as

$$\begin{aligned}
X^2_{Total} &= n(\hat{\mathbf{p}}_{ABC} - \mathbf{p}_C \otimes \mathbf{p}_B \otimes \mathbf{p}_A)'\mathbf{M}_{ABC}(\hat{\mathbf{p}}_{ABC} - \mathbf{p}_C \otimes \mathbf{p}_B \otimes \mathbf{p}_A)\\
&= n(\hat{\mathbf{p}}'_{ABC}\mathbf{M}_{ABC}\hat{\mathbf{p}}_{ABC} - 1)\\
&= n\hat{\mathbf{p}}'_{ABC}\mathbf{M}^{1/2}_{ABC}[\mathbf{I}_{ABC} - (\mathbf{R}_{1/C} \otimes \mathbf{R}_{1/B} \otimes \mathbf{R}_{1/A})]\\
&\qquad\qquad\qquad\qquad\qquad\qquad\times\mathbf{M}^{1/2}_{ABC}\hat{\mathbf{p}}_{ABC},\\
&= n\hat{\mathbf{p}}'_{ABC}\mathbf{M}^{1/2}_{ABC}\mathbf{Q}_{1/ABC}\mathbf{M}^{1/2}_{ABC}\hat{\mathbf{p}}_{ABC},
\end{aligned} \tag{24}$$

where $\mathbf{Q}_{1/ABC} = \mathbf{I}_{ABC} - \mathbf{R}_{1/ABC}$ and $\mathbf{R}_{1/ABC} = \mathbf{R}_{1/C} \otimes \mathbf{R}_{1/B} \otimes \mathbf{R}_{1/A}$. The $\mathbf{R}_{1/ABC}$ and $\mathbf{Q}_{1/ABC}$ are both orthogonal projectors having similar properties to $\mathbf{R}_{1/AB}$ and $\mathbf{Q}_{1/AB}$ defined in the previous section. Also, the n times squared norm of the first of the seven terms as

$$X_A^2 = n\hat{\mathbf{p}}'_{ABC}\mathbf{M}_{ABC}^{1/2}(\mathbf{R}_{1/C} \otimes \mathbf{R}_{1/B} \otimes \mathbf{R}_{1/A})\mathbf{M}_{ABC}^{1/2}\hat{\mathbf{p}}_{ABC}. \tag{25}$$

The X_B^2 and X_C^2 can be defined analogously to Equation (25). Furthermore, the n times squared norm of the fifth term in (23) is equal to

$$X_{AB}^2 = n\hat{\mathbf{p}}'_{ABC}\mathbf{M}_{ABC}^{1/2}(\mathbf{R}_{1/C} \otimes \mathbf{Q}_{1/B} \otimes \mathbf{Q}_{1/A})\mathbf{M}_{ABC}^{1/2}\hat{\mathbf{p}}_{ABC}. \tag{26}$$

The X_{AC}^2 and X_{BC}^2 can be defined analogously to Equation (26). Finally, the n times squared norm of the last term of (23) is given by

$$X_{ABC}^2 = n\hat{\mathbf{p}}'_{ABC}\mathbf{M}_{ABC}^{1/2}(\mathbf{Q}_{1/C} \otimes \mathbf{Q}_{1/B} \otimes \mathbf{Q}_{1/A})\mathbf{M}_{ABC}^{1/2}\hat{\mathbf{p}}_{ABC}. \tag{27}$$

We then have the following partition of $\mathbf{X}_{Total}^2$:

$$X_{Total}^2 = X_A^2 + X_B^2 + X_C^2 + X_{AB}^2 + X_{AC}^2 + X_{BC}^2 + X_{ABC}^2. \tag{28}$$

The first three terms on the righthand side of Equation (28) pertain to the main effects of the three variables, the next three terms to the two-way interactions, and the last one to the three-way interaction effect.

Let $p_{a..}$, $p_{.b.}$, and $p_{..c}$ denote the a^{th}, b^{th}, and c^{th} element of $\mathbf{p}_A$, $\mathbf{p}_B$, and $\mathbf{p}_C$, respectively, and let $\hat{p}_{a..}$, $\hat{p}_{.b.}$, and $\hat{p}_{..c}$ represent their observed counterparts. Let $\hat{p}_{abc}$ denote the observed joint proportion of the event abc. Further, let $\hat{p}_{ab.}$, $\hat{p}_{a.c}$, and $\hat{p}_{.bc}$ represent the observed two-way marginal probabilities derived from $\hat{p}_{abc}$'s by summing them up over the subscript replaced by a dot (e.g., $\hat{p}_{ab.} = \sum_{c=1}^{C} \hat{p}_{abc}$). Then, the total chi-square in Equation (24) and the terms of its partition (from Equations (25) through (27)) can be re-expressed, in scalar notation, as

$$X_{Total}^2 = n\sum_{a=1}^{A}\sum_{b=1}^{B}\sum_{c=1}^{C}\frac{1}{p_{a..}p_{.b.}p_{..c}}(\hat{p}_{abc} - p_{a..}p_{.b.}p_{..c})^2, \tag{29}$$

$$X_A^2 = n\sum_{a=1}^{A}\frac{1}{p_{a..}}(\hat{p}_{a..} - p_{a..})^2, \tag{30}$$

$$X_{AB}^2 = n\sum_{a=1}^{A}\sum_{b=1}^{B}\frac{1}{p_{a..}p_{.b.}}(\hat{p}_{ab.} - \hat{p}_{a..}p_{.b.} - p_{a..}\hat{p}_{.b.} + p_{a..}p_{.b.})^2, \tag{31}$$

and

$$\begin{aligned} X_{ABC}^2 = n\sum_{a=1}^{A}\sum_{b=1}^{B}\sum_{c=1}^{C}\frac{1}{p_{a..}p_{.b.}p_{..c}}&(\hat{p}_{abc} - \hat{p}_{ab.}p_{..c} - \hat{p}_{a.c}p_{.b.} - \hat{p}_{.bc}p_{a..} \\ &+\hat{p}_{a..}p_{.b.}p_{..c} + \hat{p}_{.b.}p_{a..}p_{..c} + \hat{p}_{..c}p_{a..}p_{.b.} - p_{a..}p_{.b.}p_{..c})^2. \end{aligned} \tag{32}$$

The X_B^2 and X_C^2 can be re-expressed similarly to Equation (30), and X_{AC}^2 and X_{BC}^2 can be re-expressed similarly to Equation (31), by permuting the subscripts appropriately. The correspondence between matrix and scalar representations can be easily established similarly to the two-way case. Note that Equations (26) and (27) (and Equations (31) and (32)) give explicit expressions of X_{AB}^2 and X_{ABC}^2, which were defined by

$$X_{AB}^2 = n \sum_{a=1}^{A} \sum_{b=1}^{B} \frac{1}{p_{a..}p_{.b.}} (\hat{p}_{ab.} - p_{a..}p_{.b.})^2 - X_A^2 - X_B^2,$$

and

$$X_{ABC}^2 = X_{Total}^2 - X_A^2 - X_B^2 - X_C^2 + X_{AB}^2 + X_{AC}^2 + X_{BC}^2$$

in Lancaster (1951).

2.5 Extensions to higher-order tables

It is now straightforward to extend the above derivations to higher-order tables. For four-way tables,

$$\begin{aligned}
&\mathbf{M}_{ABCD}^{1/2}\hat{\mathbf{p}}_{ABCD} \\
&\quad = [(\mathbf{R}_{1/D} + \mathbf{Q}_{1/D}) \otimes (\mathbf{R}_{1/C} + \mathbf{Q}_{1/C}) \otimes (\mathbf{R}_{1/B} + \mathbf{Q}_{1/B}) \\
&\qquad\qquad \otimes (\mathbf{R}_{1/A} + \mathbf{Q}_{1/A})]\mathbf{M}_{ABCD}^{1/2}\hat{\mathbf{p}}_{ABCD}, \qquad (33)
\end{aligned}$$

where $\mathbf{M}_{ABCD} = \mathbf{D}_D^{-1} \otimes \mathbf{D}_C^{-1} \otimes \mathbf{D}_B^{-1} \otimes \mathbf{D}_A^{-1}$, and $\hat{\mathbf{p}}_{ABCD}$ is the vector of observed probabilities arranged in a similar way to the three-way case (i.e., in such a way that the category index for variable D is the slowest moving index). The righthand side of Equation (33) may be further elaborated similarly to the last equation in (23) to obtain a sixteen-term partition. Dropping the first term (pertaining to the grand mean), the total chi-square is partitioned into the sum of fifteen part chi-squares, four main effects, six two-way interaction effects, four three-way interaction effects, and one four-way interaction effect.

For higher-order tables, simply define, for any newly added variable Y, $\mathbf{M}_{ABCD\cdots Y} = \mathbf{D}_Y^{-1} \otimes \mathbf{M}_{ABCD\cdots}$, add $(\mathbf{R}_{1/Y} + \mathbf{Q}_{1/Y})\otimes$ immediately following the left bracket on the righthand side of Equation (33), and construct $\hat{\mathbf{p}}_{ABCD\cdots Y}$ in such a way that the category index for variable Y is the slowest moving index. In the general case of m variables, there will be 2^m terms in the familywise partition of the total chi-square. Dropping the first term (pertaining to the grand mean), the number of the j^{th} order interaction effects can be calculated by $\begin{pmatrix} m \\ j \end{pmatrix} = \frac{m!}{j!(m-j)!}$.

3 Scenarios concerning the specification of marginal probabilities

In deriving the partitions of the total chi-squares in the previous section, the problem of specifying marginal probabilities ($\mathbf{p}_A$, $\mathbf{p}_B$, etc.) was left open. There are two conceivable scenarios for the specification of marginal probabilities for each variable. One is in which marginal probabilities are prescribed by the analyst. The prescribed values, denoted as $\mathbf{p}_A^*$, are supposed to come from sources independent from the data. They may be theoretically driven or derived on the basis of prior knowledge about the phenomena of concern. In rare cases, they may be set equal across all categories of a variable, with a lack of strong theory or prior knowledge that suggest otherwise. The variable whose marginal probabilities are prescribed is called the *Scenario 1 variable* hereafter. Often, however, there are no such theories or prior knowledge, and marginal probabilities have to be estimated from the data. The variable whose marginal probabilities are estimated (i.e., $\mathbf{p}_A = \hat{\mathbf{p}}_A$) is called the *Scenario 2 variable*. Note that the variable whose margins are *a priori* fixed is also called a Scenario 2 variable. Distributional properties as well as calculated values of the total and the main effect part chi-squares depend on the profiles of scenarios of the variables concerned. Distributional properties of the part chi-squares for the interaction effects, on the other hand, remain intact, although their values change. In the following subsections, we elaborate on these points for tables of different orders and under different profiles of scenarios.

3.1 One-way tables

In one-way tables, the variable of concern (variable A) must be a Scenario 1 variable (because X^2_{Total} is identically equal to 0 otherwise). Then, $X^2_{Total} = X^2_A$ in Equation (9) asymptotically follows the chi-square distribution with $A-1$ df (χ^2_{A-1}) if the prescribed value of $\mathbf{p}_A = \mathbf{p}_A^*$ is correct. To see this, it is useful to rely on some standard result on the asymptotic distribution of a quadratic form involving central normal variables. Specifically, let $\mathbf{x}$ asymptotically follows a central normal distribution with the covariance matrix $\boldsymbol{\Sigma}$ (i.e., $\mathbf{x} \rightsquigarrow \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$). Then, the quadratic form $\mathbf{x}'\mathbf{C}\mathbf{x}$ asymptotically follows the chi-square distribution with df equal to rank($\boldsymbol{\Sigma}\mathbf{C}$) if and only if

$$\boldsymbol{\Sigma}\mathbf{C}\boldsymbol{\Sigma}\mathbf{C}\boldsymbol{\Sigma} = \boldsymbol{\Sigma}\mathbf{C}\boldsymbol{\Sigma} \tag{34}$$

for further details, seeAgresti (2002, p.589) Rao (1973, p.188). In the present case, $\mathbf{x} = \sqrt{n}\mathbf{D}_A^{-1/2}(\hat{\mathbf{p}}_A - \mathbf{p}_A) \rightsquigarrow \mathcal{N}(\mathbf{0}, \mathbf{Q}_{1/A})$ and $\mathbf{C} = \mathbf{Q}_{1/A}$, so that Equation (34) is satisfied. The df of this chi-square is rank$(\mathbf{Q}_{1/A}) = A-1$.

3.2 Two-way tables

The variables in two-way tables can be either a Scenario 1 variable or a Scenario 2 variable. Both can be Scenario 1 variables (Profile 1), Scenario 2 variables

(Profile 2), or a mixed case in which one is a Scenario 1 variable while the other is Scenario 2 (Profile 3).

In Profile 1, X^2_{Total} in Equation (13) asymptotically follows χ^2_{AB-1} under the assumptions that the two variables are statistically independent and that the prescribed marginal probabilities are correct (see Equation (19)). In this case, we have $\mathbf{x}'\mathbf{C}\mathbf{x}$, where $\mathbf{x} = \sqrt{n}\mathbf{Q}_{1/AB}\mathbf{M}^{1/2}_{AB}\hat{\mathbf{p}}_{AB} \rightsquigarrow \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} = \mathbf{Q}_{1/AB}$ and $\mathbf{C} = \mathbf{Q}_{1/AB}$, so that $\boldsymbol{\Sigma}\mathbf{C} = \mathbf{Q}_{1/AB}$, and it can be easily seen that the condition of Equation (34) is satisfied. The df for this chi-square is $\text{rank}(\mathbf{Q}_{1/AB}) = AB - 1$. The X^2_A in Equation (16), on the other hand, asymptotically follows χ^2_{A-1} when the specified value of $\mathbf{p}_A = \mathbf{p}^*_A$ is correct (see Equation (20)), since $\mathbf{x} = \sqrt{n}\mathbf{Q}_{1/AB}\mathbf{M}^{1/2}_{AB}\hat{\mathbf{p}}_{AB} \rightsquigarrow \mathcal{N}(\mathbf{0}, \mathbf{R}_{1/B} \otimes \mathbf{Q}_{1/A})$ and $\mathbf{C} = \mathbf{R}_{1/B} \otimes \mathbf{Q}_{1/A}$, so that $\boldsymbol{\Sigma}\mathbf{C} = \mathbf{R}_{1/B} \otimes \mathbf{Q}_{1/A}$. Again, it can be readily verified that Condition (34) is satisfied. The df for this chi-square is $\text{rank}(\mathbf{R}_{1/B} \otimes \mathbf{Q}_{1/A}) = A - 1$. Similarly, $\mathbf{X}^2_B$ in Equation (17) asymptotically follows χ^2_{B-1} when the prescribed value of $\mathbf{p}_B = \mathbf{p}^*_B$ is correct (see Equation (21)), and X^2_{AB} in Equation (18) asymptotically follows $\chi^2_{(A-1)(B-1)}$ when the prescribed value of $\mathbf{p}_{AB} = \mathbf{p}^*_B \otimes \hat{\mathbf{p}}_A + \hat{\mathbf{p}}_B \otimes \mathbf{p}^*_A - \mathbf{p}^*_B \otimes \mathbf{p}^*_A$ is correct (see Equation (22)).

It remains to be seen that the (asymptotic) distributions of X^2_A, X^2_B, and X^2_{AB} are mutually independent. This can be shown as follows: Two quadratic forms $\mathbf{x}'\mathbf{C}_1\mathbf{x}$ and $\mathbf{x}'\mathbf{C}_2\mathbf{x}$ involving the same $\mathbf{x}$, where $\mathbf{x} \rightsquigarrow \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ (asymptotically) follow independent chi-square distributions with respective df if and only if

$$\boldsymbol{\Sigma}\mathbf{C}_1\boldsymbol{\Sigma}\mathbf{C}_2\boldsymbol{\Sigma} = \mathbf{O} \tag{35}$$

(Ogasawara and Takahashi, 1951). If we apply this condition to show the asymptotic independence of X^2_A and X^2_{AB}, for example, we have $\mathbf{C}_1 = \mathbf{R}_{1/B} \otimes \mathbf{Q}_{1/A}$, $\mathbf{C}_2 = \mathbf{Q}_{1/B} \otimes \mathbf{Q}_{1/A}$, and $\boldsymbol{\Sigma} = \mathbf{Q}_{1/AB}$, so that $\boldsymbol{\Sigma}\mathbf{C}_1 = \mathbf{C}_1$ and $\boldsymbol{\Sigma}\mathbf{C}_2 = \mathbf{C}_2$. Since $\mathbf{C}_1\mathbf{C}_2 = \mathbf{O}_{AB}$ (see Equation (14)), then Equation (35) is satisfied. Thus, the independence issue reduces to the orthogonality issue in the present case. Independence of the asymptotic distributions of other pairs of part chi-squares can be similarly verified.

In Profile 2, the chi-square variables above are redefined using the estimates of $\mathbf{p}_A$ and $\mathbf{p}_B$ instead of their prescribed values. Let $\hat{\mathbf{R}}_{1/A} = \hat{\mathbf{D}}^{1/2}_A\mathbf{1}_A\mathbf{1}_A\hat{\mathbf{D}}^{1/2}_A$ and $\hat{\mathbf{Q}}_{1/A} = \mathbf{I}_A - \hat{\mathbf{R}}_{1/A}$. Further, define $\hat{\mathbf{R}}_{1/B}$ and $\hat{\mathbf{Q}}_{1/B}$ analogously. Then, (11) can be rewritten as

$$\begin{aligned}\hat{\mathbf{D}}^{-1/2}_A\hat{\mathbf{P}}_{AB}\hat{\mathbf{D}}^{-1/2}_B &= \hat{\mathbf{R}}_{1/A}\hat{\mathbf{D}}^{-1/2}_A\hat{\mathbf{P}}_{AB}\hat{\mathbf{D}}^{-1/2}_B\hat{\mathbf{R}}_{1/B} \\ &+\hat{\mathbf{Q}}_{1/A}\hat{\mathbf{D}}^{-1/2}_A\hat{\mathbf{P}}_{AB}\hat{\mathbf{D}}^{-1/2}_B\hat{\mathbf{R}}_{1/B} + \hat{\mathbf{R}}_{1/A}\hat{\mathbf{D}}^{-1/2}_A\hat{\mathbf{P}}_{AB}\hat{\mathbf{D}}^{-1/2}_B\hat{\mathbf{Q}}_{1/B} \\ &+\hat{\mathbf{Q}}_{1/A}\hat{\mathbf{D}}^{-1/2}_A\hat{\mathbf{P}}_{AB}\hat{\mathbf{D}}^{-1/2}_B\hat{\mathbf{Q}}_{1/B}.\end{aligned} \tag{36}$$

The second term on the righthand side of the above partition is

$$\begin{aligned}&\hat{\mathbf{Q}}_{1/A}\hat{\mathbf{D}}^{-1/2}_A\hat{\mathbf{P}}_{AB}\hat{\mathbf{D}}^{-1/2}_B\hat{\mathbf{R}}_{1/B} \\ &= (\mathbf{I}_A - \hat{\mathbf{D}}^{1/2}_A\mathbf{1}_A\mathbf{1}'_A\hat{\mathbf{D}}^{1/2}_A)\hat{\mathbf{D}}^{-1/2}_A\hat{\mathbf{P}}_{AB}\hat{\mathbf{D}}^{-1/2}_B\hat{\mathbf{D}}^{1/2}_B\mathbf{1}_B\mathbf{1}'_B\hat{\mathbf{D}}^{1/2}_B \\ &= \hat{\mathbf{D}}^{-1/2}_A\hat{\mathbf{P}}_{AB}\mathbf{1}_B\mathbf{1}'_B\hat{\mathbf{D}}^{1/2}_B - \hat{\mathbf{D}}^{1/2}_A\mathbf{1}_A\mathbf{1}'_A\hat{\mathbf{P}}_{AB}\mathbf{1}_B\mathbf{1}'_B\hat{\mathbf{D}}^{1/2}_B = \mathbf{O}_{AB},\end{aligned}$$

where $\mathbf{O}_{AB}$ is the zero matrix of order $A \times B$, so that X_A^2 is identically equal to 0 with 0 df. Similarly, $X_B^2 = 0$. The fourth term, on the other hand, is equal to

$$\begin{aligned}
\hat{\mathbf{Q}}'_{1/A}\hat{\mathbf{D}}_A^{-1/2}\hat{\mathbf{P}}_{AB}\hat{\mathbf{D}}_B^{-1/2}(\mathbf{I}_B - \hat{\mathbf{R}}_{1/B}) \\
&= \hat{\mathbf{Q}}_{1/A}\hat{\mathbf{D}}_A^{-1/2}\hat{\mathbf{P}}_{AB}\hat{\mathbf{D}}_B^{-1/2} \\
&= \hat{\mathbf{D}}_A^{-1/2}(\hat{\mathbf{P}}_{AB} - \hat{\mathbf{D}}_A\mathbf{1}_A\mathbf{1}'_B\hat{\mathbf{D}}_B)\hat{\mathbf{D}}_B^{-1/2},
\end{aligned} \tag{37}$$

so that

$$\begin{aligned}
X_{AB}^2 &= n\text{vec}(\hat{\mathbf{P}}_{AB} - \hat{\mathbf{D}}_A\mathbf{1}_A\mathbf{1}'_B\hat{\mathbf{D}}_B)'(\hat{\mathbf{D}}_B^{-1} \otimes \hat{\mathbf{D}}_A^{-1})\text{vec}(\hat{\mathbf{P}}_{AB} - \hat{\mathbf{D}}_A\mathbf{1}_A\mathbf{1}'_B\hat{\mathbf{D}}_B) \\
&= n(\hat{\mathbf{p}}_{AB} - \hat{\mathbf{p}}_B \otimes \hat{\mathbf{p}}_A)'(\hat{\mathbf{D}}_B^{-1} \otimes \hat{\mathbf{D}}_A^{-1})(\hat{\mathbf{p}}_{AB} - \hat{\mathbf{p}}_B \otimes \hat{\mathbf{p}}_A).
\end{aligned} \tag{38}$$

Since $\hat{\mathbf{D}}_A$ and $\hat{\mathbf{D}}_B$ approach true values of $\mathbf{D}_A$ and $\mathbf{D}_B$, respectively as n goes to infinity, the above X_{AB}^2 asymptotically follows $\chi^2_{(A-1)(B-1)}$ when the two variables are statistically independent. Note that in this case $X_{Total}^2 = X_{AB}^2$. As noted earlier, the distributional property of X_{AB}^2 remains the same as in Profile 1, although its realized (observed) value typically differs. The above chi-square term can be written, in scalar notation, as

$$X_{Total}^2 = X_{AB}^2 = n\sum_{a=1}^{A}\sum_{b=1}^{B}\frac{1}{\hat{p}_{a.}\hat{p}_{.b}}(\hat{p}_{ab} - \hat{p}_{a.}\hat{p}_{.b})^2, \tag{39}$$

which can also be obtained from Equation (22) by replacing both $p_{a.}$ and $p_{.b}$ by their estimates (i.e., $\hat{p}_{a.} = \sum_b \hat{p}_{ab}$ and $\hat{p}_{.b} = \sum_a \hat{p}_{ab}$).

In Profile 3, in which one variable (say, variable A) is under Scenario 2, while the other is under Scenario 1, we set $\mathbf{p}_A = \hat{\mathbf{p}}_A$ and $\mathbf{p}_B = \mathbf{p}_B^*$. The X_A^2 becomes identically equal to zero, while X_B^2 remains the same as in Profile 1. The X_{AB}^2 in Equation (22) becomes

$$X_{AB}^2 = n\sum_{a=1}^{A}\sum_{b=1}^{B}\frac{1}{\hat{p}_{a.}p^*_{.b}}(\hat{p}_{ab} - \hat{p}_{a.}\hat{p}_{.b})^2, \tag{40}$$

where $p^*_{.b}$ is the b^{th} element of $\mathbf{p}_B^*$. The X_{AB}^2 above asymptotically follows the $\chi^2_{(A-1)(B-1)}$ under the assumptions that the two variables are statistically independent, and that the prescribed marginal probabilities for variable B are correct. Again, the distributional property remains unchanged, although its realised value calculated from the data is likely to be different from that calculated under different profiles. The X_{Total}^2 asymptotically follows $\chi^2_{A(B-1)}$ under the same conditions as the asymptotic null distribution of X_{AB}^2 is obtained.

3.3 Three-way tables

Cases of three-way tables are similar. There are cases in which all three variables are under Scenario 1 (Profile 1), cases in which all three variables are under Scenario 2 (Profile 2), and mixed cases in which some variables are under Scenario 1 while the others are under Scenario 2 (Profile 3), although there are more variety of profiles in Profile 3 than in two-way tables.

In Profile 1, we set $\mathbf{p}_Y = \mathbf{p}_Y^*$ for all variables Y = A, B, and C in X^2_{Total}, X^2_A, X^2_B, X^2_C, X^2_{AB}, X^2_{AC}, X^2_{BC}, and X^2_{ABC} defined in Section 2.3. They all asymptotically follow the chi-square distributions with $ABC-1$, $A-1$, $B-1$, $C-1$, $(A-1)(B-1)$, $(A-1)(C-1)$, $(B-1)(C-1)$, and $(A-1)(B-1)(C-1)$ df, respectively, all under the hypotheses that the prescribed marginal probabilities are correct, and that all three variables are statistically independent.

In Profile 2, all three main effect chi-squares (X^2_A, X^2_B andX^2_C) become identically 0 with 0 df, so that

$$X^2_{Total} = X^2_{AB} + X^2_{AC} + X^2_{BC} + X^2_{ABC}.$$

All three two-way interaction part-chi-squares become similar to Equation (39) with an additional dot in the subscripts replacing the index of the variable not involved in the interaction (e.g., we replace $\hat{p}_{ab}$ by $\hat{p}_{ab.}$, $\hat{p}_{a.}$ by $\hat{p}_{a..}$, $p_{.b}$ by $p_{.b.}$ in case of X^2_{AB}). The two-way interaction chi-squares, X^2_{AB}, X^2_{AC}, and X^2_{BC}, all asymptotically follow the chi-square distributions with $(A-1)(B-1)$, $(A-1)(C-1)$, and $(B-1)(C-1)$ df, respectively, under the hypothesis that all three variables are statistically independent. The three-way interaction chi-square in Equation (32) becomes

$$X^2_{ABC} = n\sum_{a=1}^{A}\sum_{b=1}^{B}\sum_{c=1}^{C}\frac{1}{\hat{p}_{a..}\hat{p}_{.b.}\hat{p}_{..c}}(\hat{p}_{abc}-\hat{p}_{ab.}\hat{p}_{..c}-\hat{p}_{a.c}\hat{p}_{.b.}-\hat{p}_{.bc}\hat{p}_{a..}+2\hat{p}_{a..}\hat{p}_{.b.}\hat{p}_{..c})^2, \tag{41}$$

which asymptotically follows $\chi^2_{(A-1)(B-1)(C-1)}$ under the hypothesis that all three variables are independent. The total chi-square, X^2_{Total}, asymptotically follows $\chi^2_{ABC-A-B-C+2}$ under the same condition.

In Profile 3, only those main effect chi-squares corresponding to the Scenario 2 variables become 0 with 0 df, while those under Scenario 1 are unaffected. There are three possible cases in the two-way interaction chi-squares, say X^2_{AB}. If both A and B are the Scenario 1 variables (Case 1), Equation (31) remains essentially unchanged. If A is under Scenario 2, and B is under Scenario 1 (Case 2), X^2_{AB} will be similar to Equation (40). If both A and B are under Scenario 2 (Case 3), X^2_{AB} will be similar to Equation (39). In all cases, X^2_{AB} asymptotically follows $\chi^2_{(A-1)(B-1)}$ under the null hypothesis that depends on the cases. In Case 1, the hypothesis is that all three variables are independent and that the prescribed marginal probabilities are correct for both variables. In Case 2, it is that the three variables are independent and that the prescribed marginal probabilities for variable B are correct. In Case 3, only the independence among the three variables is required. The three-way interaction chi-square is obtained by replacing the vectors of hypothesised marginal probabilities in Equation (32) by their observed counterparts only for variables under Scenario 2. This is a mixture of Equations (32) and (41). In all cases, X^2_{ABC} asymptotically follows $\chi^2_{(A-1)(B-1)(C-1)}$ under the hypothesis that the three variables are statistically independent, and that the prescribed marginal probabilities are correct for all Scenario 1 variables. Under the same conditions, X^2_{Total} asymptotically follows the chi-square distribution with df equal to $ABC - 1$ minus the df's of the Scenario 2 variables.

4 Illustrative examples

In this section, we present two exemplary analyses of three-way contingency tables, based on the partitions of Pearson's chi-square statistic described in Section 2 under the scenarios in the specification of marginal probabilities discussed in Section 3. As emphasized earlier, one benefit of our general partitions of the chi-square statistic is that it is valid whether the marginal probabilities are prescribed (Scenario 1) or estimated from the data or fixed *a priori* (Scenario 2). The former allows us to perform simultaneous tests of marginal probabilities and independence among the variables. Often, only one of these is conducted, ignoring the other.

4.1 An example data from Lang (1996)

The first data set we analyze is a contingency table from Lang (1996, p.1021) displayed in Table 1. There are three variables (A, B, and C) with two categories (1 and 2) in each. Following Lang (1996), all three variables are treated as the Scenario 1 variables with prescribed probabilities of $\mathbf{p}^*_A = (0.5, 0.5)'$, $\mathbf{p}^*_B = (0.5, 0.5)'$, and $\mathbf{p}^*_C = (0.5, 0.5)'$ (marginal homogeneity across all three variables). The part chi-squares in the seven-term partition of the total chi-square derived under the assumptions of complete marginal homogeneity and three-way independence are displayed in the second column of Table 2 along with df's and p-values in columns 3 and 4. These values were obtained using the function *chi2scen1* of the R package *chi2x3way* (Lombardo et al., 2017). Note that our partition is exact, contrary to Lang's suggested partition. As the tests of interactions all depend on the correctness of prescribed marginal probabilities, we first look at the values of the main effect chi-squares, none of which turned out to be significant at the 5% level. This indicates that the hypothesized marginal homogeneity is reasonably correct for all three variables, which implies that the interaction chi-squares are assured to give unconfounded tests of interactions from mis-specifications of the marginal probabilities. Only the B by C interaction has turned out to be significant at the 5% level. This implies that $\hat{p}_{a..} \approx p^*_{a..} = 1/2$, $\hat{p}_{.b.} \approx p^*_{.b.} = 1/2$, $\hat{p}_{..c} \approx p^*_{..c} = 1/2$, $\hat{p}_{ab.} \approx p_{a..}p_{.b.}$, $\hat{p}_{a.c} \approx p_{a..}p_{..c}$, and $X^2_{ABC} \approx 0$ (where "$\approx$" means "approximately equal"), which in turn implies $\hat{p}_{abc} \approx \hat{p}_{.bc}/2$ for all a, b, and c. This suggests that the model in which one variable (variable A) is independent from the other two (B and C) with equal probabilities of categories of variable A is the most plausible model for this data set.

Table 1 The three-way contingency table from (Lang, 1996, p.1021).

		C	
A	B	1	2
1	1	5	10
	2	11	3
2	1	9	8
	2	9	4

Table 2 Partition of Pearson's three-way chi-square statistic (X^2) under Scenario 1 for all three variables and three alternative partitions of the log-likelihood ratio chi-square (G^2), obtained from the data in Table 1.

				G^2		
effect	X^2	df	p-value	Partition 1	Partition 2	Partition 3
A	0.017	1	0.896	0.017	0.017	0.017
B	0.424	1	0.515	0.424	0.424	0.424
C	1.373	1	0.241	1.378	1.378	1.378
AB	0.153	1	0.696	0.145	0.145	
AC	0.153	1	0.696	0.141		0.141
BC	4.898	1	0.027		5.650	5.650
AB|AC,BC						0.272
AC|AB,BC					0.268	
BC|AB,AC				5.777		
AB,AC,BC	5.204	3	0.158	6.063	6.063	6.063
ABC	1.373	1	0.241	1.295	1.295	1.295
AB,AC,BC,ABC	6.577	4	0.160	7.358	7.358	7.358
Total(+A,B,C)	8.390	7	0.299	9.178	9.178	9.178

As promised in the introduction section, we compare our partition of Pearson's chi-square statistic and analogous partitions of the log-likelihood ratio chi-square (G^2), often used in log-linear analyses of contingency tables. Unfortunately, the familywise partition of G^2_{Total} is not unique for tables of order higher than two. Depending on how one splits the joint effects of three two-way interaction effects, three alternative partitions emerge (Cheng et al., 2006; Loisel and Takane, 2016). Specifically, let G^2_{AB} denote the part G^2 due to the AB interaction ignoring the other two two-way interaction effects, and let $G^2_{AB|AC,BC}$ (where the subscripts AB are followed by AC, BC separated by $|$) denote the part G^2 due to the AB interaction eliminating the effects of AC and BC interactions (Goodman (1969) called the former the marginal two-way interaction between A and B, and the latter the partial two-way interaction between A and B after eliminating the effects of AC and BC interactions). Then, $G^2_{AB,AC,BC} = G^2_{AB} + G^2_{AC} + G^2_{BC|AB,AC} = G^2_{AB} + G^2_{BC} + G^2_{AC|AB,BC} =$

$G^2_{AC} + G^2_{BC} + G^2_{AB|AC,BC}$. Depending on which of these three partitions of $G^2_{AB,AC,BC}$ we use, three different partitions of G^2_{Total} result. Columns 5 to 7 of Table 2 report the part G^2's for the three alternative partitions of G^2_{Total} for the same data set as above. These values were obtained by *Proc Hiloglinear* and *Proc Loglinear* in SPSS. Each partition is exact, however, as can be seen from the table, unlike the partition reported in Table 2 of Lang (1996). Note that G^2_{ABC} can only be calculated iteratively. Since the tests of partial two-way interactions depend on the assumption of no three-way interaction effect (Goodman, 1969), we start with the test of the three-way interaction, which happens to be nonsignificant. We then proceed to the tests of partial two-way interaction effects. As it happens, only the BC interaction effect is significant. If we retain this effect in the model, the marginal BC interaction is no longer testable. The only marginal two-way interactions that can be tested independently from the partial BC interaction are the AB and AC interactions, neither of which turn out to be significant. Since the BC interaction effect is significant, the only main effect that can be meaningfully tested is the A main effect (Andersen, 1980, 1991), which turns out to be nonsignificant. The whole pattern of the test results indicates that the model, in which variable A is independent from the other two with the additional marginal homogeneity of categories of variable A, is the best model. Thus, essentially the same conclusion as above reached by Pearson's chi-square statistics can be drawn using the LR statistics (Andersen, 1980, p. 104). See Loisel and Takane (2016) for more detailed comparisons between partitions of Pearson's statistics and the LR ratio statistics.

It seems difficult, if not impossible, to make general remarks on relative merits of X^2 and G^2. There are a variety of criteria (e.g., speed of convergence, efficiency, robustness against outliers, overdispersion, etc.) and a wide range of conditions under which their performance is evaluated. Results are "mixed" depending on the conditions and criteria. For example, focussing on efficiency alone, the two statistics are asymptotically equivalent in one situation (fixed cells, a fixed size of the tests, and local alternatives). On the other hand, G^2 is said to be optimal in another situation (fixed cells, fixed (nonlocal) alternatives, and size of the tests tending to zero as the sample size increases), while this superiority of G^2 does not necessarily hold, if the number of cells is assumed to increase at the same rate as the sample size (Hoeffding, 1965). Thus, there is no statistic which is universally the best. One promising idea to resolve the issue is the use of the power-divergence statistic proposed by Cressie and Read (1984), which is a family of statistics indexed by the value of λ, and which includes both G^2 ($\lambda = 0$) and X^2 ($\lambda = 1$) as its special cases. Cressie and Read (1984) recommend $\lambda = 2/3$ as a good compromise between G^2 and X^2; see also Read and Cressie (1988, p. 96-97). While it is not easy to derive exact partitions of the power-divergence statistic with $\lambda = 2/3$ similar to the ones for X^2, it is possible to turn each term in the partitions of X^2 into a power-divergence statistic by plugging in appropriate values of observed and expected joint or marginal probabilities in the formula of the power-divergence statistic.

4.2 Twin births by year, gender combinations, and race

Here we illustrate the partitions of Pearson's chi-square statistic in a genetic study. The data concerns the number of twin births in the U.S. Birth Registration Area from 1922 to 1936 (inclusive) annually for Caucasian and African-American twins of the same and different gender (Strandskov and Edelen, 1946). Table 3 displays a contingency table of dimension $15 \times 3 \times 2$, consisting of the following three variables: Variable A is the birth year with 15 categories, variable B represents the gender profiles in twins with three categories, both males (MM), both females (FF), and one male and one female (MF), and variable C is the race of twins with two categories, Caucasian and African-American.

Table 3 Crosstabulation of twins in terms of birth year (variable A), combination of gender (variable B), and race (variable C).

(C)Race	Caucasian			African-American		
(A)Year\(B)Gender	MM	FF	MF	MM	FF	MF
1922	6176	6304	6467	735	794	687
1923	6298	6298	6547	735	815	751
1924	6552	6659	7052	775	916	797
1925	6412	6173	6875	651	746	674
1926	6412	6309	6864	697	747	697
1927	7334	7357	7998	1056	1082	1033
1928	7499	7555	7998	1164	1323	1247
1929	7422	7504	7721	1251	1319	1272
1930	7224	7316	7583	1289	1432	1284
1931	6782	7074	7239	1280	1432	1259
1932	6790	7102	7164	1253	1468	1304
1933	6721	6926	7094	1377	1545	1327
1934	7135	7054	7401	1388	1558	1457
1935	6911	6983	7385	1211	1414	1293
1936	7104	7059	7492	1290	1329	1295

There is a substantive theory deducing the marginal probabilities of different gender profiles. Human twins are generally thought to be of two types; monozygotic (MZ) (or single-egg) twins and dizygotic (DZ) (or two-egg) twins. In human genetics, the existence of dizygotic twins explains why some twins consist of different gender. That monozygotic twins occur at all is attested by the fact that more same-sex twins appear than what one would expect if the twins were all dizygotic. For the genetic rules on twinning, we may assume that the probability of twins being dizygotic is $P(DZ) = 2/3$ and the probability of being monozygotic is $P(MZ) = 1/3$. Also, let the probability of two twin males and two twin females being born, given that they are monozygotic, be

$P(MM|MZ) = P(FF|MZ) = 1/2$, while the probability of different genders in monozygotic twins be $P(MF|MZ) = 0$. Further, the probabilities of two male dizygotic twins, and two female dizygotic twins are $P(MM|DZ) = 1/4$ and $P(FF|DZ) = 1/4$, respectively, while $P(MF|DZ) = 1/2$. This leads to $P(MM) = 1/3, P(FF) = 1/3$, and $P(MF) = 1/3$, so that the prescribed marginal probabilities for the gender variable are $\mathbf{p}_B^* = (1/3, 1/3, 1/3)'$. In the first analysis of the above data set, the gender profile variable (variable B) is regarded as a Scenario 1 variable so we are going to test that its probabilities are homogeneous, while the other two as Scenario 2 variables whose marginal probabilities are estimated from the data.

Table 4 summarises the part chi-squares in the partition of Pearson's statistic, given the above specifications of the marginal probabilities. The partition includes one nonzero main effect term (X_B^2), three two-way interaction terms, and one three-way interaction term. The table reports, for each term, their magnitude, the corresponding df, and the p-value under the assumptions of complete independence among the three variables and of the correctness of the prescribed marginal probabilities for variable B. The total chi-square is $X_{Total}^2 = 2362.66$ with 74 df, which is highly significant, indicating a clear departure from independence and/or mis-specifications of marginal probabilities. Note, however, that X_{Total}^2 does not tell us which of the familywise effects are responsible for its significance. For this, we need to look at the values of part chi-squares in the partition of this quantity.

We first consider the main effect of variable B ($X_B^2 = 159.39$ with 2 df), which is highly significant, so we reject the hypothesis that the marginal probabilities are homogeneous indicating a "gross" mis-specifications of the marginal probabilities of this variable. There must be something wrong with the theory from which the expected marginal probabilities were derived. It may be that the assumption of $P(MZ) = 1/3$ and $P(DZ) = 2/3$ and/or that of $P(M) = P(F) = 1/2$ from which conditional probabilities of gender profiles given the types of twins are deduced is too crude. In any case, this result indicates that there is a room for improvements in the genetic theory.

The genetic theory given above also did not take into account the possibility that the values of the probabilities used to calculate the expected marginal probabilities of the twin's gender profiles may depend on other variables such as the year of birth and the race of the twin-pair that is born. In order to check this possibility, we look at other part chi-squares in the table. The three-way interaction ($X_{ABC}^2 = 23.66$ with 28 df) is not significant, but three two-way interactions are all significant in varying degrees. In particular, the gender by race interaction ($X_{BC}^2 = 149.8$ with 2 df) is highly significant, indicating the probabilities of the three gender profiles (MM, FF, and MF) do differ across the two races. The gender by year interaction is also significant ($X_{AB}^2 = 54.03$ with 28 df), although much less significant than the gender by race interaction. This indicates that the probabilities of the gender profiles also change as the years go by (although the nature of the changes is unclear at this point). The year by race interaction is by far the largest interaction effect ($X_{AC}^2 = 1975.99$ with 14 df) among all the interaction effects. This indicates that the probability distribution of twin births across the two races (regardless of their gender profiles) change in time (although this effect has nothing to do with the genetic theory that predicted the probabilities of gender profiles).

Table 4 Partition of Pearson's three-way chi-square statistic obtained from the data in Table 3 under Scenario 1 for variable B (gender profile) and under Scenario 2 for variables A (year) and C (race).

effect	X^2	df	p-value
B	159.39	2	<0.001
AB	54.03	28	0.002
AC	1975.99	14	<0.001
BC	149.58	2	<0.001
ABC	23.66	28	0.699
Total	2362.66	74	<0.001

The analyses reported above assumed that the effects of the mis-specification of marginal probabilities for gender profiles on interaction chi-squares were negligible, despite the fact that the main effect of the gender variable was significant. The interaction chi-squares related to variable B reported in Table 4 may be un-ignorably affected by the mis-specification. To find it out, the same data set was reanalysed, assuming that all three variables were under Scenario 2. Table 5 reports the recalculated values of the interaction chi-squares. As it turns out, all part chi-square values are very similar to the corresponding part chi-squares in the previous table. Apparently, the effects of the mis-specification are relatively minor in this particular instance. This may be because the observed marginal probabilities (of 0.325, 0.332, and .342) are not far away from the prescribed value of .333 for all three gender profiles, with the significance of the departures being primarily driven by the large sample size ($n = 365,774$). A big question is how we know when the mis-specification of the marginal probabilities significantly affect the values of interaction chi-squares, and when not? Unfortunately, there is no good answer to this question *a priori*. The only way that can be recommended is, as in the present case, to reanalyse the data regarding the variables as under Scenario 2, whose main effects are found to be significant under Scenario 1, and compare the results.

Table 5 Partition of Pearson's three-way chi-square statistic obtained from the data in Table 3 under Scenario 2 for all three variables.

Effect	$\mathbf{X}^2$	df	p-value
AB	53.85	28	0.002
AC	1976.00	14	<0.001
BC	148.50	2	<0.001
ABC	22.73	28	0.746
Total	2201.08	72	<0.001

5 Discussion

In this paper, we presented a general framework for partitioning Pearson's chi-square statistic under the assumption of complete independence among the variables. This framework is "general", as it applies to contingency tables of any order. The resulting partitions are unique, exact, and can be calculated in closed form, unlike some of their competitors (e.g., the log-likelihood ratio chi-square). With these partitions, simultaneous tests of marginal and joint probabilities in contingency tables become feasible. Although these tests demand an additional burden of prescribing marginal probabilities, they, in return, provide opportunities to kill "two birds with one stone", whenever plausible values of marginal probabilities are available. In case they are not available, one can always resort back to the usual analyses of the joint probabilities, ignoring the marginals. Two examples of analyses along this line were presented to illustrate the point.

Plackett (1962) has claimed that X^2_{ABC} in Lancaster (1951) (Equation (27) in this paper) does not need to follow an asymptotic chi-square distribution under the notion of no three-way interaction effect originally proposed by Bartlett (1935); see also Simpson (1951) and Roy and Kastenbaum (1956). This is indeed true if it is the only plausible definition of no three-way interaction. We argue that there is an alternative definition, more suitable in the context of Pearson's chi-square statistic, namely

$$p_{abc} = \hat{p}_{ab.}p_{..c} + \hat{p}_{a.c}p_{.b.} + \hat{p}_{.bc}p_{a..} - \hat{p}_{a..}p_{.b.}p_{..c} - \hat{p}_{.b.}p_{a..}p_{..c} \\ - \hat{p}_{..c}p_{a..}p_{.b.} + p_{a..}p_{.b.}p_{..c} \qquad (42)$$

for all combinations of a, b, and c (see Equation (32)). Under this notion of no three-way interaction, our theoretical contention that the asymptotic null distribution of X^2_{ABC} is well approximated by a chi-square distribution under a variety of conditions (profiles of scenarios) has been confirmed by extensive Monte-Carlo studies (Lombardo et al., 2017).

Note that familywise partitions of the total chi-squares into multiple part chi-squares tend to increase the number of tests to be performed on one data set. For example, there could be up to seven tests in three-way contingency tables. To keep the joint α level (the probability of committing a Type 1 error in at least one of the tests performed) to a reasonable level, we may use the Bonferroni type of tests, as has been recommended by Andersen (1980, 1991). In these tests, the usual α level is divided by the number of tests to be performed, and each test is performed with the reduced α level. This guarantees the joint α level is at most the prescribed value of α. A potential drawback of the Bonferroni tests is that they are often too conservative. In fact, they get more and more conservative as the number of tests to be performed. Another interesting idea is to use post-hoc tests like the ones developed by Goodman (1964), which keep the joint α level constant no matter how many tests are performed.

The partitions of Pearson's statistic presented in this paper were all derived under the assumption of complete independence among the variables. Some argue that this is too restrictive because this assumption is violated too often in practical situations. One may ask, however, how we will know when the assumption is violated? To find it out, we need to know the null distribution of the statistic (X^2_{Total}), and to identify the nature of the violation, that of the terms in the partition of the statistic (part chi-squares). The assumption of independence is essential as a working hypothesis to derive such distributions.

If the purpose of the analyses goes beyond the complete independence assumption, we can offer alternative partitions of Pearson's statistic under weaker assumptions, such as one variable independent from other two (one-factor independence), or two variables conditionally independent given a third variable (conditional independence). In fact, such partitions were already derived and used by Loisel and Takane (2016, Section 4.3) in the analysis of three-way contingency tables. It should be noted, however, that there are three ways in which one variable is independent from the other two, and also three ways in which two variables are conditionally independent from a third variable in three-way contingency tables, and these numbers tend to go up rapidly, as the tables increase their order. As an appropriate partition depends on these situational factors, we have to be aware in advance which situation we are in. This could be a bit too demanding in many practical situations.

Acknowledgements On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Agresti, A. (1990). *Categorical Data Analysis.* John Wiley & Sons, New York, USA.

Agresti, A. (2002). *Categorical Data Analysis, Second Ed.* John Wiley & Sons, New York, USA.

Andersen, E. B. (1980). *The Statistical Analysis of Categorical Data.* Springer-Verlag.

Andersen, E. B. (1991). *The Statistical Analysis of Categorical Data (Second, Revised and Enlarged Edition).* Springer-Verlag.

Bartlett, M. S. (1935). Partitioning pearson's chi-squared statistic for a partially ordered three-way contingency table. *Journal of the Royal Statistical Society, Series B*, 2:248–252.

Beh, E. J. and Davy, P. J. (1998). Partitioning pearson's chi-squared statistic for a completely ordered three-way contingency table. *The Australian and New Zealand Journal of Statistics*, 40:465–477.

Beh, E. J. and Lombardo, R. (2014). *Correspondence Analysis, Theory, Practice and New Strategies.* John Wiley & Sons, Chichester, UK.

Beh, E. J., Simonetti, B., and D'Ambra, L. (2007). Partitioning a non-symmetric measure of association for three-way contingency tables. *Journal of Multivariate Analysis*, 98:1391–1411.

Bishop, Y. M. M., Fienberg, S. E., and Holland, P. W. (1975). *Discrete Multivariate Analysis.* MA: MIT Press, Cambridge, UK.

Carlier, A. and Kroonenberg, P. M. (1996). Decompositions and biplots in three-way correspondence analysis. *Psychometrika*, 61:355–373.

Cheng, P. E., Liou, J. W., Liou, M., and Aston, A. D. (2006). Data information in contingency tables: A fallacy of hierarchical loglinear models. *Journal of Data Science*, 4:387–398.

Cressie, N. and Read, T. C. R. (1984). Multinomial goodness-of-fit tests. *Journal of the Royal Statistical Society, Series B*, 46:440–464.

Friendly, M. (1994). Mosaic displays for multi-way contingency tables. *Journal of the American Statistical Association*, 89:190–200.

Goodman, L. A. (1964). Simultaneous confidence intervals for contrasts among multinomial populations. *The Annals of Mathematical Statistics*, 35:716–725.

Goodman, L. A. (1969). On partitioning χ^2 and detecting partial association in three-way contingency tables. *Journal of the Royal Statistical Society, Series B*, 31:486–498.

Goodman, L. A. (1970). The multivariate analysis of qualitative data: Interactions among multiple classifications. *Journal of the American Statistical Associations*, 65:226–256.

Hoeffding, W. (1965). Asymptotically optimal tests for multinomial distributions. *The Annals of Mathematical Statistics*, 36:369–408.

Kroonenberg, P. M. (2008). *Applied Multiway Data Analysis.* John Wiley & Sons, Hoboken, NJ.

Lancaster, H. O. (1951). Complex contingency tables treated by the partition of the chi-square. *Journal of Royal Statistical Society, Series B*, 13:242–249.

Lang, J. B. (1996). On the partitioning of goodness-of-fit statistics for multivariate categorical response models. *Journal of the American Statistical Association*, 91:1017–1023.

Loisel, S. and Takane, Y. (2016). Partitions of pearson's chi-square statistic for frequency tables: A comprehensive account. *Computational Statistics*, 31:1429–1452.

Lombardo, R., Carlier, A., and D'Ambra, L. (1996). Nonsymmetric correspondence analysis for three-way contingency tables. *Methodologica*, 4:59–80.

Lombardo, R., Takane, Y., and Beh, E. (2017). Chi2x3way package. https://cran.r-project.org/web/packages/chi2x3way/index.html.

Ogasawara, T. and Takahashi, T. (1951). Independence of quadratic forms of a random sample from a normal population. *Science Bulletin, Hiroshima University*, 15:1–9.

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. *Philosophical Magazine*, 1:157–175.

Plackett, R. L. (1962). A note on interactions in contingency tables. *Journal of the Royal Statistical Society, Series B*, 24:162–166.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications.* John Wiley & Sons, New-York, USA.

Read, T. C. R. and Cressie, N. (1988). *Goodness-of Fit Statistics for Discrete Multivariate Data.* Springer-Verlag, New-York, USA.

Roy, S. N. and Kastenbaum, M. A. (1956). On the hypothesis of no "interaction" in a multi-way contingency table. *Annals of Mathematical Statistics*, 27:749–757.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Series B, 13, 238-241*, 13:238–241.

Strandskov, H. H. and Edelen, W. E. (1946). Monozygotic and dizygotic twin birth frequencies in the total, the white and the colored us population. *Genetics*, 31:438–446.

van der Heijden, P. G. M. and de Leeuw, J. (1985). Correspondence analysis used complementary to loglinear analysis. *Psychometrika*, 50:429–447.

Yanai, H., Takeuchi, K., and Takane, Y. (2011). *Projection Matrices, Generalized Inverse Matrices, and Singular Value Decomposition.* Springer, New-York, USA.