AISM (2007) 59: 557–576
DOI 10.1007/s10463-006-0058-4

P. Vellaisamy · V. Vijay

Some collapsibility results for *n*-dimensional contingency tables

Received: 8 December 2004 / Revised: 26 December 2005 /
Published online: 26 April 2007

Abstract For a multidimensional contingency table, we obtain several necessary and sufficient conditions for collapsibility and strict collapsibility, using the technique of Möbius inversion formula. As a consequence, the results of Whittemore (*Journal of the Royal Statistical Society B, 40*, 328–340, 1978) are stated in a form which is easy to understand and the proofs are much simpler and straightforward. Several new results on collapsibility and strict collapsibility with respect to more than one interaction parameter, are established, and their relationships to conditional independence are also pointed out. As applications of our results, several typical examples on collapsibility, strict collapsibility and conditional independence are discussed. It is also shown that Bishop et al. (*Discrete Multivariate Analysis: Theory and Practice*, MIT Press, Cambridge, 1975) conditions are necessary and sufficient for strict collapsibility with respect to a set of interaction factors.

Keywords Collapsibility · Conditional independence · Contingency table · Log-linear model · Möbius inversion · Simpson's paradox · Strict collapsibility

1 Introduction

The analysis of a large dimensional contingency table, using a log-linear model or other approach is quite involved. It is often very helpful to reduce the dimension of the table or convenient to look at the condensed (summed over certain variables) table. In a condensed table, some extraneous association between the remaining variables may be introduced or an original relationship between certain variables may be lost and/or the monotonicity of certain dependence between

P. Vellaisamy (✉) · V. Vijay
Department of Mathematics, Indian Institute of Technology,
Bombay, Powai, Mumbai 400 076, India
E-mail: pv@math.iitd.ac.in

variables may also be reversed. This paradox, due to Simpson (1951), known as Simpson's paradox. See, also, Lindley and Novick (1981) and Shapiro (1982) and Cox and Wermuth (2003) for more examples. However, there are certain tables which do not exhibit Simpson's paradox. In such cases, it is advantageous to condense (more technically, collapse) the original table, especially when the observed counts are small in many cells of the table (Ducharme and Lepage 1986). Hence, it is of practical importance to identify tables which are collapsible. In view of recent advances in information technology, there is a huge amount of data available, and the problems of data mining have become statistically challenging ones. Hence, collapsibility may also be viewed as a "dimension reduction problem" or condensation of the data. For example, Wermuth (1987) studied the parametric collapsibility with respect to odds ratio and relative risk, and Guo and Geng (1995) discussed the collapsibility conditions for logistic regression coefficients. Whittemore (1978) obtained, for a n-dimensional table (that is, the data on n-categorical variables), some necessary and sufficient conditions for collapsibility and strict collapsibility. However, the results are rather difficult to understand, as they are based on certain functional representations. The proofs of some of the results are also difficult, because of the algebraic approach adopted there. Our collapsibility results are based on interaction factors, appearing in the log-linear models, rather than in terms of certain arbitrary functions. In this paper, we adopt a novel approach based on Möbius inversion formula, which is a well-known technique in combinatorial methods (Charalambides, 2002). Note that the use of Möbius inversion is not new in statistics. In fact, it is used as an essential step for proving Hammersley–Clifford theorem (Lauritzen, 1996, p. 36 or Hammersley and Clifford 1971) which is fundamental for the whole field of graphical models.

In Sect. 2, the log-linear model and some basic results on interaction factors are presented. In Sect. 3, we obtain a set of equivalent conditions for collapsibility. As a consequence, Whittemore's (1978) necessary and sufficient conditions for collapsibility are stated in a compact form involving interaction factors, and the proofs are also simple and straightforward. Even for moderately large dimensional contingency tables, it is of practical importance to collapse the table with respect to more than one interaction factor. There are no such results available in the literature. We obtain necessary and sufficient conditions for collapsibility with respect to τ_L, where $L \subseteq A$ (with $|A| = r$), and $k \in L$, and also for the case $\{j, k\} \in L$. As applications, suitable examples are also illustrated. In Sect. 4, we obtain similar results for strict collapsibility, and bring out their relationship to conditional independence of two variables, given other variables, for an n-dimensional table. Using this result, we show that Bishop et al. (1975) conditions are necessary and sufficient for strict collapsibility with respect to a set of parameters.

2 The log-linear model for an n-dimensional table

We start with the following result, called Möbius inversion theorem (Lauritzen 1996, p. 239) which is used as an important tool for analyzing the log-linear models, interaction factors and proving some results on collapsibility. For more details on posets and properties of Möbius functions, the reader is referred to Charalambides (2002, p. 167).

Lemma 2.1 *Let f and g be functions defined on a locally finite poset P containing a zero element 0. Then*

$$f(x) = \sum_{y \leq x} g(y), \quad \forall x \in P, \tag{1}$$

if and only if

$$g(x) = \sum_{y \leq x} \mu(y, x) f(y), \quad \forall x \in P,$$

where $\mu(x, y)$ is the Möbius function. If f and g satisfy Eq. (1)*, then we call (f, g) a Möbius pair.*

Remark 2.1 Let P be a finite poset with the partial order relation $\subseteq$ (is a subset of) and be denoted by $(P, \subseteq)$. In that case, the Möbius function is,

$$\mu(A, B) = \begin{cases} 1, & \text{if } A = B \\ (-1)^{|B-A|}, & \text{if } A \subset B \\ 0, & \text{if } A \not\subseteq B. \end{cases}$$

Let $\bar{n} = \{1, 2, \ldots, n\}$. Let $X_1, \ldots, X_n$ be n categorical variables with the support $S(X_j) = \bar{m}_j$, for $1 \leq j \leq n$. Let $i = (i_1, \ldots, i_n)$ denote a cell of the n-dimensional table, where $1 \leq i_k \leq m_k$. Let $p(i)$ be the probability that the individual observation $X = (X_1, \ldots, X_n)$ falls in the i-th cell, that is, $p(i) = p(i_1, \ldots, i_n) = P(X_1 = i_1, \cdots, X_n = i_n)$. Assume $p(i) > 0$ and $\sum p(i) = 1$. Define $l(i) = \ln p(i)$. For example, when $n = 3$ (3-dimensional table), the log-linear model is defined as (see, Bishop et al., 1975, p. 33

$$\begin{aligned} l^{(3)}(i_1, i_2, i_3) &= \tau_{123}^{(3)}(i_1, i_2, i_3) + \tau_{12}^{(3)}(i_1, i_2) + \tau_{13}^{(3)}(i_1, i_3) + \tau_{23}^{(3)}(i_2, i_3) \\ &\quad + \tau_1^{(3)}(i_1) + \tau_2^{(3)}(i_2) + \tau_3^{(3)}(i_3) + \tau_\phi^{(3)} \\ &= \sum_A \tau_A^{(3)}(i_A), \end{aligned}$$

where A is any subset of $\{1, 2, 3\}$. Also, it can be seen that, for example

$$\begin{aligned} \tilde{l}_{12}^{(3)}(i_1, i_2) &:= \frac{1}{m_3} \sum_{i_3} l^{(3)}(i_1, i_2, i_3), \\ &= \tau_{12}^{(3)}(i_1, i_2) + \tau_1^{(3)}(i_1) + \tau_2^{(3)}(i_2) + \tau_\phi^{(3)} \\ &= \sum_{Z \subseteq \{1,2\}} \tau_Z^{(3)}(i_Z). \end{aligned}$$

Consider next the n-dimensional contingency table with

$$l^{(n)}(i) = l(i_1, \cdots, i_n) = \ln(p(i_1, \cdots, i_n)).$$

Let

$$l^{(n)}(i) = \sum_{Z \subseteq \bar{n}} \tau_Z^{(n)}(i_Z)$$

be the log-linear model, where $\tau_A^{(n)}(i_A)$ denotes the r-factor interaction, if $|A|=r$. For example, when $\bar{n} = \{1, 2, \ldots, 10\}$, and $A = \{1, 3, 5\}$, then $\tau_A^{(n)}(i_A) = \tau_{135}^{(10)}(i_1, i_3, i_5)$ denotes the three factor interaction between X_1, X_3 and X_5. In fact, our aim is to first obtain $\tau_Z^{(n)}(i_Z)$ and establish its properties. Define

$$\tilde{l}_A^{(n)}(i_A) = \frac{1}{\prod_{j \in A^c} m_j} \sum\nolimits_{i_j : j \in A^c} l^{(n)}(i). \tag{2}$$

Then, as observed in three-dimensional table,

$$\tilde{l}_A^{(n)}(i_A) = \sum_{Z \subseteq A} \tau_Z^{(n)}(i_Z), \quad \forall A \subseteq \bar{n}. \tag{3}$$

Indeed, we will prove this result rigorously later (see Lemma 2.3), after obtaining the expression for $\tau_Z(i_Z)$.

Considering the poset $(\mathcal{P}, \subseteq)$, and $\tilde{l}_A$ as the function on $(\mathcal{P}, \subseteq)$, we obtain from Eq. (3), using Möbius inversion theorem (see Remark 2.1),

$$\tau_A^{(n)}(i_A) = \sum_{Z \subseteq A} \mu(Z, A) \tilde{l}_Z^{(n)}(i_Z) = \sum_{Z \subseteq A} (-1)^{|A-Z|} \tilde{l}_Z^{(n)}(i_Z), \quad \forall A \subseteq \bar{n}. \tag{4}$$

Note that by Möbius inversion theorem (see Lemma 2.1), Eq. (3) holds if and only if Eq. (4) holds. Observe also that Möbius inversion can not be applied directly to $l^{(n)}(i)$ to obtain the form of $\tau_A^{(n)}$.

Remark 2.2 Whittemore (1978) defined first $\tau_A^{(n)}(i_A)$, given in Eq. (4), as a straightforward extension, and later remarked that $l(i) = \sum_{Z \subseteq \bar{n}} \tau_Z(i_Z)$. However, our approach is more direct, simpler and brings out various hidden properties of the interaction factors and log-linear models.

We next establish some basic properties of τ_A's.

Lemma 2.2 *Let $A = \{1, \ldots, r\}$. Then, the interaction parameters $\tau_A^{(n)}$, defined in Eq.* (4)*, satisfies*

$$\sum_{i_k} \tau_A^{(n)}(i_A) = 0, \quad \forall k \in A.$$

Proof Let $A_k = A \setminus \{k\}$, and Z_k be any subset of A_k. From Eq. (4),

$$\begin{aligned}
\sum_{i_k} \tau_A^{(n)}(i_A) &= \sum_{i_k} \sum_{Z \subseteq A} (-1)^{|A-Z|} \tilde{l}_Z(i_Z) \\
&= \sum_{i_k} \left[\sum_{Z \subseteq A; k \in Z} (-1)^{|A-Z|} \tilde{l}_Z(i_Z) + \sum_{Z \subseteq A; k \notin Z} (-1)^{|A-Z|} \tilde{l}_Z(i_Z) \right] \\
&= m_k \left\{ \sum_{Z \subseteq A; k \in A} (-1)^{|A-Z|} \left(\frac{1}{m_k} \sum_{i_k} \tilde{l}_Z(i_Z) \right) + \sum_{Z \subseteq A; k \notin Z} (-1)^{|A-Z|} \left(\frac{1}{m_k} \sum_{i_k} \tilde{l}_Z(i_Z) \right) \right\} \\
&= m_k \left\{ \sum_{Z_k \subseteq A_k} (-1)^{|A_k-(Z_k \cup k)|} \tilde{l}_{Z_k}(i_{Z_k}) + \sum_{Z_k \subseteq A_k} (-1)^{|A_k - Z_k|} \tilde{l}_{Z_k}(i_{Z_k}) \right\} \\
&= m_k \left\{ \sum_{Z_k \subseteq A_k} \left[(-1)^{|A_k-(Z_k \cup k)|} + (-1)^{|A_k - Z_k|} \right] \tilde{l}_{Z_k}(i_{Z_k}) \right\} \\
&= 0,
\end{aligned}$$

which completes the proof. □

Note that the above result is true for any interaction factor $\tau_A^{(n)}$, where A is any subset of $\bar{n}$.

Remark 2.3 Whittemore (1978) first proved the above result for an n-dimensional table. However, her proof is complicated as it uses her Lemma 3.1 which is rather involved. Our proof is simpler, straightforward, and uses the only fact $\{Z | Z \subseteq A; k \in Z\} = \{Z_k \cup \{k\} | Z_k \subseteq A_k\}$.

The next lemma shows another interesting property of log-linear model.

Lemma 2.3 *Let $\tau_Z(i_Z)$, $Z \subseteq \bar{n}$, satisfy Lemma* 2.2*, Then, the log-linear model*

$$l^{(n)}(i) = \sum_{Z \subseteq \bar{n}} \tau_Z^{(n)}(i_Z) \tag{5}$$

if and only if

$$\tilde{l}_A(i_A) = \sum_{Z \subseteq A} \tau_Z^{(n)}(i_Z), \quad \forall A \subseteq \bar{n}, \tag{6}$$

where $\tilde{l}_A^{(n)}$ is defined in Eq. (2).

Proof Let

$$l^{(n)}(i) = \sum_{Z \subseteq \bar{n}} \tau_Z^{(n)}(i_Z),$$

where the $\tau_Z^{(n)}$ satisfy Lemma 2.2, and $m_A = \prod_{j\in A^c} m_j$. From Eq. (2),

$$\begin{aligned}
\tilde{l}_A^{(n)}(i) &= \frac{1}{m_A}\sum_{i_j:j\in A^c} l^{(n)}(i),\\
&= \frac{1}{m_A}\sum_{i_j:j\in A^c}\sum_{Z\subseteq \bar{n}} \tau_Z^{(n)}(i_Z)\\
&= \frac{1}{m_A}\left[\sum_{i_j:j\in A^c}\left\{\sum_{Z:Z\cap A^c=\phi}\tau_Z^{(n)}(i_Z) + \sum_{Z:Z\cap A^c\neq\phi}\tau_Z^{(n)}(i_Z)\right\}\right]\\
&= \sum_{Z\subseteq A}\tau_Z^{(n)}(i_Z) + \frac{1}{m_A}\sum_{Z:Z\cap A^c\neq\phi}\left\{\sum_{i_j:j\in Z^c\cap A^c}\sum_{i_j:j\in Z\cap A^c}\tau_Z^{(n)}(i_Z)\right\}\\
&= \sum_{Z\subseteq A}\tau_Z^{(n)}(i_Z),
\end{aligned}$$

since $\sum_{i_j:j\in Z\cap A^c}\tau_Z^{(n)}(i_Z) = 0$, by Lemma 2.2. The sufficiency part trivially follows by taking $A = \bar{n}$. □

The above lemma justifies our approach of assuming Eq. (3) and obtaining the representation for $\tau_A^{(n)}$, given in Eq. (4), using the Möbius inversion formula.

3 Collapsibility

In this section, we obtain some necessary and sufficient conditions for collapsibility for an n-dimensional contingency table. First we introduce the basic notations.

Let $A = \{a_1, \ldots, a_r\}$, and $B = \{a_1, \ldots, a_r, a_{r+1}, \ldots, a_s\}$, where $|A| = r \leq s = |B| < n$, be any two subsets of $\bar{n}$. Let now

$$p_B(i_B) = \sum_{i_j:j\in B^c} p(i)$$

denote the cell probabilities of the marginal (condensed over B^c) table. Define, similarly,

$$l^{(s)}(i) = \ln(p_B(i_B))$$

and

$$\tilde{l}_Z^{(s)}(i_Z) = \frac{1}{\prod_{j\in B\setminus Z} m_j}\sum_{i_j:j\in B\setminus Z} l^{(s)}(i), \tag{7}$$

where $Z \subseteq B$. Let

$$l^{(s)}(i) = \sum_{Z\subseteq B}\eta_Z^{(s)}(i_Z)$$

be the log-linear model for the marginal table. It follows from Lemma 2.3 that,

$$\tilde{l}_A^{(s)}(i_A) = \sum_{Z \subseteq A} \eta_Z^{(s)}(i_Z) \tag{8}$$

for every $A \subseteq B$. Note in general $\eta_Z^{(s)} \neq \tau_Z^{(n)}$, defined in Eq. (5).

The following definition is due to Whittemore (1978) and is sometimes also called model-collapsibility.

Definition 3.1 (Collapsibility) *An n-dimensional table is* collapsible *(over B^c) to an s-dimensional table with respect to $\tau_A^{(n)}$, $A \subseteq B$, if*

$$\tau_A^{(n)}(i_A) = \eta_A^{(s)}(i_A) \quad \textit{for all} \;\; i_A.$$

Let now

$$d^{(B)}(i_B) = \ln(p_B(i_B)) - \tilde{l}_B^{(n)}(i_B) = l^{(s)}(i) - \tilde{l}_B^{(n)}(i_B) \tag{9}$$

and for any $Z \subseteq B$

$$\tilde{d}_Z^{(B)}(i_Z) = \frac{1}{\prod_{j \in B \setminus Z} m_j} \sum_{i_j : j \in B \setminus Z} d^{(B)}(i_B). \tag{10}$$

The following result characterizes the situations under which collapsibility holds with respect to $\tau_A^{(n)}$, $A \subseteq B$.

Theorem 3.1 *Let $l^{(n)}(i) = \sum_{Z \subseteq \bar{n}} \tau_Z^{(n)}(i_Z)$, and $l^{(s)}(i) = \sum_{Z \subseteq B} \eta_Z^{(s)}(i_Z)$ be respectively the log-linear models for an n-dimensional and the marginal s-dimensional tables. Let $\delta_Z = (\eta_Z^{(s)} - \tau_Z^{(n)})$, for $Z \subseteq B$. Then the following conditions are equivalent:*

(i) $\delta_A(i_A) = 0$;
(ii) $\tilde{d}_A^{(B)}(i_A) = \sum_{Z \subset A} \delta_Z(i_Z)$;
(iii) $\sum_{Z \subseteq A} (-1)^{|A-Z|} \tilde{d}_Z^{(B)}(i_Z) = 0$,

where $\tilde{d}_Z^{(B)}$ is defined in Eq. (10), *and $A \subseteq B$.*

Proof From Eq. (9),

$$l^{(s)}(i) = \tilde{l}_B^{(n)}(i_B) + d^{(B)}(i_B). \tag{11}$$

Also, for any $Z \subseteq B$, we have from Eqs. (7), (10) and (11)

$$\begin{aligned}\tilde{l}_Z^{(s)}(i_Z) &= \frac{1}{\prod_{j \in B \setminus Z} m_j} \sum_{i_j : j \in B \setminus Z} \left\{ \tilde{l}_B^{(n)}(i_B) + d^{(B)}(i_B) \right\} \\ &= \tilde{l}_Z^{(n)}(i_Z) + \tilde{d}_Z^{(B)}(i_Z)\end{aligned}$$

which, using Eqs. (6) and (8), leads to

$$\tilde{d}_Z^{(B)}(i_Z) = \tilde{l}_Z^{(s)}(i_Z) - \tilde{l}_Z^{(n)}(i_Z) \tag{12}$$
$$= \sum_{X \subseteq Z} (\eta_X^{(s)}(i_X) - \tau_X^{(n)}(i_X))$$
$$= \sum_{X \subseteq Z} \delta_X(i_X). \tag{13}$$

Applying Möbius inversion formula to Eq. (13), with $Z = A$,

$$\delta_A(i_A) = \sum_{Z \subseteq A} (-1)^{|A-Z|} \tilde{d}_Z^{(B)}(i_Z). \tag{14}$$

Thus, (i)$\Longleftrightarrow$(ii) follows from Eq. (13) and (i)$\Longleftrightarrow$(iii) follows from Eq. (14). □

Remark 3.1 The equivalence of (i) and (ii) is similar to Theorem 2 of Whittemore (1978). But, her result involves sums of certain functions which are not specifically stated. Whereas, our conditions are more explicit and involve only the interaction functions, which are easy to understand.

Next we obtain a necessary and sufficient condition for collapsibility with respect to more than one parameter.

First note that a log-linear model $l^{(n)}(i) = \sum_{Z \subseteq \bar{n}} \tau_Z^{(n)}$ is said to be hierarchical if $\tau_B^{(n)} \neq 0 \implies \tau_A^{(n)} \neq 0$ for $A \subset B$ or equivalently $\tau_C^{(n)} = 0 \implies \tau_D^{(n)} = 0$ for $D \supset C$ (See, Whittaker 1990 or Simonoff 2003, p. 319). Consider now the following table:

X_4	1				2			
X_3	1		2		1		2	
X_2								
X_1	1	2	1	2	1	2	1	2
1	30	120	2	16	16	65	1	9
2	490	400	10	55	80	65	2	9

Consider the hierarchical log-linear model [12][23][34], namely,

$$l^{(4)}(i) = \tau_\phi^{(4)} + \tau_1^{(4)}(i) + \tau_2^{(4)}(j) + \tau_3^{(4)}(k) + \tau_4^{(4)}(l) + \tau_{12}^{(4)}(i, j) + \tau_{23}^{(4)}(j, k) + \tau_{34}^{(4)}(k, l).$$

Then the expected numbers (cell counts) for the above data (Andersen, 1990, p. 233) under the model [12][23][34] are:

X_4	1				2			
X_3	1		2		1		2	
X_2								
X_1	1	2	1	2	1	2	1	2
1	30.43	119.56	1.8	16.37	16.58	65.15	.98	8.92
2	489.54	400.32	10.33	54.82	79.45	64.97	1.68	8.9

For the above table, we obtain

$$\tau_\phi = 3.1567,\ \ \tau_1(1) = 1.333,\ \ \tau_2(1) = -0.6308, \tau_3(1) = -0.5653,\ \ \tau_4(1) = 0.6064;$$
$$\tau_{12}(1, 1) = 0.339,\ \ \tau_{23}(1, 1) = -0.2638,\ \ \tau_{34}(1, 1) = -0.3028.$$

Suppose we collapse the table over X_2. Then for marginal table

$$\eta^{(3)}_{13}(1, 1) = -0.126,\quad \eta^{(3)}_{14}(1, 1) = 0,\quad \eta^{(3)}_{34}(1, 1) = -0.3028,\quad \eta^{(3)}_{4}(1) = 0.6064.$$

Thus, we see that Simpson's paradox occurs, since there is no or weak (conditional) association between X_1 and X_3 in the original table, but strong association between those variables in the marginal table. However, if one is interested in studying relationships between X_4 and other variables, (e.g., conditional independence between X_4 and X_1), then the table can be collapsed into a 3-dimensional table with respect to τ_L, where $L \subseteq \{1, 3, 4\}$ and $\{4\} \in L$.

The above example motivates the following general result.

Theorem 3.2 *Let $B \subset \bar{n}$ with $|B| = s < n$, and $A \subseteq B$. An n-dimensional table is collapsible (over B^c) into s-dimensional table with respect to the set $C_k = \{\tau^{(n)}_L | \{k\} \subseteq L \subseteq A\}$ of interaction factors if and only if*

$$\tilde{d}^{(B)}_A(i_A) = \tilde{d}^{(B)}_{A_k}(i_{A_k}), \tag{15}$$

where $A_k = A \setminus \{k\}$ and $1 \leq k \leq s$.

Proof Using Eq. (12), the condition

$$\tilde{d}^{(B)}_{A_k}(i_{A_k}) = \tilde{d}^{(B)}_A(i_A)$$

is equivalent to

$$\tilde{l}^{(n)}_A(i_A) - \tilde{l}^{(n)}_{A_k}(i_{A_k}) = \tilde{l}^{(s)}_A(i_A) - \tilde{l}^{(s)}_{A_k}(i_{A_k}),$$

as A and A_k are subsets of B. Using Eqs. (6) and (8), the above equation is equivalent to

$$\sum_{Z\subseteq A} \tau^{(n)}_Z(i_Z) - \sum_{Z\subseteq A_k} \tau^{(n)}_Z(i_Z) = \sum_{Z\subseteq A} \eta^{(s)}_Z(i_Z) - \sum_{Z\subseteq A_k} \eta^{(s)}_Z(i_Z),$$

or, equivalently,

$$\sum_{Z\subseteq A_k} \tau^{(n)}_{Z\cup k}(i_Z, i_k) = \sum_{Z\subseteq A_k} \eta^{(s)}_{Z\cup k}(i_Z, i_k). \tag{16}$$

We next show that Eq. (16) holds if and only if the n-dimensional table is collapsible with respect to $\tau^{(n)}_L$, for any $L \subseteq A$ and $k \in L$.

First suppose Eq. (16) holds. Summing over all i_j except i_k, we obtain

$$\tau^{(n)}_k(i_k) = \eta^{(s)}_k(i_k), \tag{17}$$

since, by Lemma 2.2, $\sum_{i_j} \tau_{Z\cup\{k\}}(i_Z, i_k) = 0$ for every nonempty set $Z \subseteq A_k$. Substituting Eq. (17) in Eq. (16), we get

$$\sum_{\substack{Z\subseteq A_k\\ Z\neq\phi}} \tau^{(n)}_{Z\cup k}(i_Z, i_k) = \sum_{\substack{Z\subseteq A_k\\ Z\neq\phi}} \eta^{(s)}_{Z\cup k}(i_Z, i_k). \tag{18}$$

Summing now over i_m, for all $m \in A_{jk} = A \setminus \{j, k\}$ in Eq. (18), leads to

$$\tau^{(n)}_{jk}(i_j, i_k) = \eta^{(s)}_{jk}(i_j, i_k).$$

Continuing this process, we get

$$\tau^{(n)}_{L}(i_L) = \eta^{(s)}_{L}(i_L), \quad \text{for all } L \subseteq A, \quad \text{and } k \in L.$$

Conversely, assume now the table is collapsible with respect to τ_L, so that

$$\tau^{(n)}_{L}(i_L) = \eta^{(s)}_{L}(i_L), \quad \text{for all} \quad i_L,$$

where $L \subseteq A$ and $k \in L$. This implies

$$\sum_{\substack{L\subseteq A\\ k\in L}} \tau^{(n)}_{L}(i_L) = \sum_{\substack{L\subseteq A\\ k\in L}} \eta^{(s)}_{L}(i_L),$$

which is equivalent to

$$\sum_{Z\subseteq A_k} \tau^{(n)}_{Z\cup k}(i_Z, i_k) = \sum_{Z\subseteq A_k} \eta^{(s)}_{Z\cup k}(i_Z, i_k)$$

and thus Eq. (16) holds. This completes the proof. □

We next consider an application of the Theorem 3.2.

Example 3.1 (Simonoff, 2003, p. 354) The following table contains the data on a clinical trial involving 1502 premature infants, conducted to study the effectiveness of the drug Palivizumab. The subjects were given either placebo or Palivizumab to see if they were hospitalized for RSV (respiratory syncytial virus). Let X_1 denotes the treatment ($X_1 = 1$ for placebo and $X_1 = 2$ for Palivizumab), $X_2 = 1$ denotes RSV hospitalization (while $X_2 = 2$, its negation) and X_3 denotes the location 1,2 and 3 respectively for US, UK, Canada.

X_3	1		2		3	
	X_2					
X_1	1	2	1	2	1	2
1	44	382	4	36	5	29
2	39	812	3	80	6	62

For the above table, it can be seen that

$$\tilde{d}_{12}(i, j) = \begin{cases} 1.70, & \text{if } i = 1, j \in \{1, 2\} \\ 1.80, & \text{if } i = 2, j \in \{1, 2\}; \end{cases} \qquad \tilde{d}_1(i) = \begin{cases} 1.70, & \text{if } i = 1 \\ 1.80, & \text{if } i = 2 \end{cases}$$

which implies

$$\tilde{d}_{12}(i, j) = \tilde{d}_1(i).$$

By Theorem 3.2, the table is collapsible with respect to $\tau_{12}^{(3)}$ and $\tau_2^{(3)}$.

Example 3.2 Consider the table (Whittemore, 1978, p. 382) given below:

X_3	1			2			3		
	X_2								
X_1	1	2	3	1	2	3	1	2	3
1	125	40	75	40	32	120	75	24	45
2	40	32	24	32	64	96	120	96	72
3	75	120	45	24	96	72	45	72	27

For the above table, we have

$$\tilde{d}_{12}(i, j) = \begin{cases} 1.2024, & \text{if } i \in \{1, 3\}, \ j = 1 \\ 1.1204, & \text{if } i \in \{1, 3\}, \ j = 2 \\ 1.1767, & \text{if } i \in \{1, 3\}, \ j = 3 \\ 1.2767, & \text{if } i = 2, \ j = 1 \\ 1.1945, & \text{if } i = 2, \ j = 2 \\ 1.1664, & \text{if } i = 2, \ j = 3; \end{cases} \tag{19}$$

$$\tilde{d}_1(i) = \begin{cases} 1.1664, & \text{if } i = 1 \\ 1.2408, & \text{if } i = 2 \\ 1.1664, & \text{if } i = 3; \end{cases} \tag{20}$$

$$\tilde{d}_2(j) = \begin{cases} 1.1664, & \text{if } j = 1 \\ 1.1449, & \text{if } j = 2 \\ 1.2015, & \text{if } j = 3. \end{cases} \tag{21}$$

It is clear that $\tilde{d}_{12}$ is neither equal to $\tilde{d}_1$ or $\tilde{d}_2$, and so the condition of Theorem 3.2 is not satisfied. So collapsibility with respect to τ_{12} and τ_1 (or τ_{12} and τ_2) does not hold simultaneously. However, note that

$$\tau_{12}^{(3)}(i, j) = \eta_{12}^{(2)}(i, j) = \begin{cases} 0.281, & \text{if } i = 1, \ j = 1 \\ -0.562, & \text{if } i = 1, \ j = 2 \\ 0.281, & \text{if } i = 1, \ j = 3 \\ -0.025, & \text{if } i = 2, \ j = 1 \\ 0.05, & \text{if } i = 2, \ j = 2 \\ -0.025, & \text{if } i = 2, \ j = 3 \\ -0.26, & \text{if } i = 3, \ j = 1 \\ 0.52, & \text{if } i = 3, \ j = 2 \\ -0.26, & \text{if } i = 3, \ j = 3 \end{cases}$$

Hence, collapsibility with respect to $\tau_{12}^{(3)}$ holds.

The above example serves as the basis for the following result.

Theorem 3.3 *Let $B \subset \bar{n}$ with $|B| = s < n$, and $A \subseteq B$. An n-dimensional table is collapsible (over B^c) into an s-dimensional table with respect to the set $C_{j,k} = \{\tau_L^{(n)} | \{j,k\} \subseteq L \subseteq A\}$ of interaction factors if and only if*

$$\tilde{d}_A^{(B)}(i_A) - \tilde{d}_{A_k}^{(B)}(i_{A_k}) = \tilde{d}_{A_j}^{(B)}(i_{A_j}) - \tilde{d}_{A_{jk}}^{(B)}(i_{A_{jk}}), \tag{22}$$

where $A_k = A \setminus \{k\}$ and $A_{jk} = A \setminus \{j,k\}$.

Proof Note first that the Eq. (22) is equivalent to

$$\begin{aligned} &\tilde{l}_A^{(n)}(i_A) - \tilde{l}_{A_k}^{(n)}(i_{A_k}) - \tilde{l}_{A_j}^{(n)}(i_{A_j}) + \tilde{l}_{A_{jk}}^{(n)}(i_{A_{jk}}) \\ &= \tilde{l}_A^{(s)}(i_A) - \tilde{l}_{A_k}^{(s)}(i_{A_k}) - \tilde{l}_{A_j}^{(s)}(i_{A_j}) + \tilde{l}_{A_{jk}}^{(s)}(i_{A_{jk}}). \end{aligned} \tag{23}$$

Using Eqs. (6) and (8), Eq. (23) becomes

$$\begin{aligned} &\sum_{Z \subseteq A} \tau_Z^{(n)}(i_Z) - \sum_{Z \subseteq A_k} \tau_Z^{(n)}(i_Z) - \sum_{Z \subseteq A_j} \tau_Z^{(n)}(i_Z) + \sum_{Z \subseteq A_{jk}} \tau_Z^{(n)}(i_Z) \\ &= \sum_{Z \subseteq A} \eta_Z^{(s)}(i_Z) - \sum_{Z \subseteq A_k} \eta_Z^{(s)}(i_Z) - \sum_{Z \subseteq A_j} \eta_Z^{(s)}(i_Z) + \sum_{Z \subseteq A_{jk}} \eta_Z^{(s)}(i_Z). \end{aligned} \tag{24}$$

Note that Eq. (24) is equivalent to

$$\sum_{Z \subseteq A_k} \tau_{Z \cup k}^{(n)}(i_Z, i_k) - \sum_{Z \subseteq A_{jk}} \tau_{Z \cup k}^{(n)}(i_Z, i_k) = \sum_{Z \subseteq A_k} \eta_{Z \cup k}^{(s)}(i_Z, i_k) - \sum_{Z \subseteq A_{jk}} \eta_{Z \cup k}^{(s)}(i_Z, i_k)$$

which in turn is equivalent to

$$\sum_{Z \subseteq A_{jk}} \tau_{Z \cup \{j,k\}}^{(n)}(i_Z, i_j, i_k) = \sum_{Z \subseteq A_{jk}} \eta_{Z \cup \{j,k\}}^{(s)}(i_Z, i_j, i_k). \tag{25}$$

Thus, (22) $\Longleftrightarrow$ (25).

Summing now over i_l, for all $l \in A_{jk}$, we get from Eq. (25),

$$\tau_{jk}^{(n)}(i_j, i_k) = \eta_{jk}^{(s)}(i_j, i_k),$$

and repeating the arguments similar to the proof of Theorem 3.2, we obtain

$$\tau_L^{(n)}(i_L) = \eta_L^{(s)}(i_L),$$

for all subsets L of A containing $\{j,k\}$.

Conversely, let

$$\tau_L^{(n)}(i_L) = \eta_L^{(s)}(i_L),$$

where $L \subseteq A$ and $j,k \in L$. This implies

$$\sum_{\substack{L \subseteq A \\ \{j,k\} \subseteq L}} \tau_L^{(n)}(i_L) = \sum_{\substack{L \subseteq A \\ \{j,k\} \subseteq L}} \eta_L^{(s)}(i_L),$$

or, equivalently

$$\sum_{Z\subseteq A_{jk}} \tau^{(n)}_{Z\cup\{j,k\}}(i_Z, i_j, i_k) = \sum_{Z\subseteq A_{jk}} \eta^{(s)}_{Z\cup\{j,k\}}(i_Z, i_j, i_k),$$

which is same as Eq. (25). This completes the proof. □

Remarks 3.1

(i) For the collapsibility of a 3-dimensional table into a 2-dimensional table with respect to $\tau^{(3)}_A$, there is only one possibility of $L = A$. Hence, Eq. (22) is necessary and sufficient condition for collapsibility with respect to $\tau^{(3)}_A$. Also, in this case, condition (22) reduces to the condition (iii) of Theorem 3.1.

(ii) Theorem 3.3 can be extended to obtain the conditions of collapsibility with respect to the set $\{\tau^{(n)}_L\}$, where L is any subset of A containing three or more indices. For example, a necessary and sufficient condition of collapsibility with respect to $\{\tau^{(n)}_L\}$, where L is any subset of A containing $\{j, k, l\}$, is

$$\tilde{d}^{(B)}_A(i_A) - \tilde{d}^{(B)}_{A_k}(i_{A_k}) = \left\{\tilde{d}^{(B)}_{A_j}(i_{A_j}) - \tilde{d}^{(B)}_{A_{jk}}(i_{A_{jk}})\right\} + \left\{\tilde{d}^{(B)}_{A_l}(i_{A_l}) - \tilde{d}^{(B)}_{A_{lk}}(i_{A_{lk}})\right\} - \left\{\tilde{d}^{(B)}_{A_{jl}}(i_{A_{jl}}) - \tilde{d}^{(B)}_{A_{jkl}}(i_{A_{jkl}})\right\}.$$

If we continue this process to the case $L = A$, the necessary and sufficient conditions for collapsibility with respect to $\tau^{(n)}_A$ (only) leads to condition (iii) of Theorem 3.1.

Example 3.2 (continued) Consider the contingency table discussed in Example 3.2. It can be shown that $\tilde{d}_\phi = 1.2$. Also, from Eqs. (19)–(21),

$$d_{12}(i, j) - \tilde{d}_1(i) = \tilde{d}_2(j) - \tilde{d}_\phi,$$

and so the condition (22) is satisfied. Hence, the table is collapsible with respect to $\tau^{(3)}_{12}$, as observed already in Example 3.2.

4 Strict collapsibility

We next look at a stronger version of collapsibility, namely, strict collapsibility (Whittemore, 1978). We assume in this section, for simplicity, $A = \{1, 2, \ldots, r\}$, and $B = \{1, 2, \ldots, s\}$, where $s \geq r$.

Definition 4.1 (Strict Collapsibility) *An n-dimensional table is said to be* strictly collapsible *into an s-dimensional table (over B) with respect to* $\tau^{(n)}_A$, $A \subseteq B$, *if*

(i) $\tau^{(n)}_A = \eta^{(s)}_A$, *and*
(ii) $\tau^{(n)}_Z = 0, \quad \forall\, Z \supseteq A, Z \cap B^c \neq \phi$.

Note, condition (i) of above definition is the definition of collapsibility. First, we obtain an equivalent condition for condition (ii) of strict collapsibility.

Lemma 4.1 *For an n-dimensional contingency table,*

$$\tau_Z^{(n)}(i_Z) = 0, \quad \textit{for all} \;\; Z \supseteq A, \;\; Z \cap B^c \neq \phi,$$

if and only if

$$\sum_{\substack{Z \supseteq A \\ Z \cap B^c \neq \phi}} \tau_Z^{(n)}(i_Z) = 0. \tag{26}$$

Proof The necessary part is obvious. Suppose now Eq. (26) holds. Then,

$$\sum_{\substack{Z \subseteq A^c \\ Z \cap B^c \neq \phi}} \tau_{A \cup Z}^{(n)}(i_A, i_Z) = 0. \tag{27}$$

Summing over $i_m, m \in A^c \setminus \{k\}$, for $k \in B^c$, leads to

$$\tau_{A \cup \{k\}}^{(n)}(i_A, i_k) = 0, \quad \forall k \in B^c. \tag{28}$$

Substituting Eq. (28) in Eq. (27), we obtain

$$\sum_{\substack{Z \subseteq A^c \\ Z \cap B^c \neq \phi, Z \neq \{k\}}} \tau_{A \cup Z}^{(n)}(i_A, i_Z) = 0, \quad \forall k \in B^c.$$

Repeating the arguments as done earlier, we get

$$\tau_Z^{(n)}(i_Z) = 0, \quad \text{for all} \;\; Z \supseteq A, \quad Z \cap B^c \neq \phi,$$

which completes the proof. □

Theorem 4.1 *Let $B \subset \bar{n}$ with $|B| = s < n$, and $A \subseteq B$. Suppose an n-dimensional contingency table is collapsible (over B^c) to an s-dimensional table with respect to $\tau_A^{(n)}$. Then, it is also strictly collapsible if and only if*

$$\sum_{Z \subseteq A} (-1)^{|A-Z|} \tilde{l}_{Z \cup A^c}^{(n)}(i_Z, i_{A^c}) = \sum_{Z \subseteq A} (-1)^{|A-Z|} \tilde{l}_{Z \cup A^c \setminus B^c}^{(n)}(i_Z, i_{A^c \setminus B^c}). \tag{29}$$

Proof First suppose Eq. (29) holds, which is equivalent to

$$\begin{aligned} l^{(n)}(i) - \tilde{l}_{A \cup A^c \setminus B^c}^{(n)}(i_A, i_{A^c \setminus B^c}) + \sum_{Z \subset A} (-1)^{|A-Z|} \Big\{ \tilde{l}_{Z \cup A^c}^{(n)}(i_Z, i_{A^c}) \\ - \tilde{l}_{Z \cup A^c \setminus B^c}^{(n)}(i_Z, i_{A^c \setminus B^c}) \Big\} = 0. \end{aligned} \tag{30}$$

Observe now that

$$\begin{aligned} l^{(n)}(i) - \tilde{l}_{A \cup A^c \setminus B^c}^{(n)}(i_A, i_{A^c \setminus B^c}) &= \sum_{Z \subseteq \bar{n}} \tau_Z^{(n)}(i_Z) - \sum_{Z \subseteq A \cup A^c \setminus B^c} \tau_Z^{(n)}(i_Z), \\ &= \sum_{\substack{Z \subseteq \bar{n} \\ Z \cap B^c \neq \phi}} \tau_Z^{(n)}(i_Z), \\ &= \sum_{\substack{Z \supseteq A \\ Z \cap B^c \neq \phi}} \tau_Z^{(n)}(i_Z) + \sum_{\substack{Z \not\supseteq A \\ Z \cap B^c \neq \phi}} \tau_Z^{(n)}(i_Z). \end{aligned} \tag{31}$$

Note also that

$$\sum_{\substack{Z\not\supseteq A\\ Z\cap B^c\neq\phi}}\tau_Z^{(n)}(i_Z)=\sum_{j=1}^{r}\sum_{\substack{Z\subseteq A_j\cup A^c\\ Z\cap B^c\neq\phi}}\tau_Z^{(n)}(i_Z)-\sum_{j,l=1}^{r}\sum_{\substack{Z\subseteq A_{jl}\cup A^c\\ Z\cap B^c\neq\phi}}\tau_Z^{(n)}(i_Z)+\cdots+(-1)^{|A|+1}\sum_{\substack{Z\subseteq A^c\\ Z\cap B^c\neq\phi}}\tau_Z^{(n)}(i_Z)$$

$$=\sum_{j=1}^{r}\left\{\sum_{Z\subseteq A_j\cup A^c}\tau_Z^{(n)}(i_Z)-\sum_{Z\subseteq A_j\cup A^c\setminus B^c}\tau_Z^{(n)}(i_Z)\right\}$$

$$-\sum_{j,l=1}^{r}\left\{\sum_{Z\subseteq A_{jl}\cup A^c}\tau_Z^{(n)}(i_Z)-\sum_{Z\subseteq A_{jl}\cup A^c\setminus B^c}\tau_Z^{(n)}(i_Z)\right\}$$

$$+\cdots+(-1)^{|A|+1}\left\{\sum_{Z\subseteq A^c}\tau_Z^{(n)}(i_Z)-\sum_{Z\subseteq A^c\setminus B^c}\tau_Z^{(n)}(i_Z)\right\}$$

$$=\sum_{j=1}^{r}\left\{\tilde{l}^{(n)}_{A_j\cup A^c}(i_{A_j},i_{A^c})-\tilde{l}^{(n)}_{A_j\cup A^c\setminus B^c}(i_{A_j},i_{A^c\setminus B^c})\right\}$$

$$-\sum_{j,l=1}^{r}\left\{\tilde{l}^{(n)}_{A_{jl}\cup A^c}(i_{A_{jl}},i_{A^c})-\tilde{l}^{(n)}_{A_{jl}\cup A^c\setminus B^c}(i_{A_{jl}},i_{A^c\setminus B^c})\right\}$$

$$+\cdots+(-1)^{|A|+1}\left\{\tilde{l}^{(n)}_{A^c}(i_{A^c})-\tilde{l}^{(n)}_{A^c\setminus B^c}(i_{A^c\setminus B^c})\right\}$$

$$=-\sum_{Z\subset A}(-1)^{|A-Z|}\left\{\tilde{l}^{(n)}_{Z\cup A^c}(i_Z,i_{A^c})-\tilde{l}^{(n)}_{Z\cup A^c\setminus B^c}(i_Z,i_{A^c\setminus B^c})\right\}. \qquad (32)$$

Substituting Eq. (32) in Eq. (31), we get

$$l^{(n)}(i)-\tilde{l}^{(n)}_{A\cup A^c\setminus B^c}(i_A,i_{A^c\setminus B^c})$$
$$=\sum_{\substack{Z\supseteq A\\ Z\cap B^c\neq\phi}}\tau_Z^{(n)}(i_Z)-\sum_{Z\subset A}(-1)^{|A-Z|}\left\{\tilde{l}^{(n)}_{Z\cup A^c}(i_Z,i_{A^c})-\tilde{l}^{(n)}_{Z\cup A^c\setminus B^c}(i_Z,i_{A^c\setminus B^c})\right\}. \qquad (33)$$

Substituting Eq. (33) in Eq. (30), we obtain

$$\sum_{\substack{Z\supseteq A\\ Z\cap B^c\neq\phi}}\tau_Z^{(n)}(i_Z)=0. \qquad (34)$$

Note all the Eqs. (29)–(34) are equivalent. The proof now follows from Lemma 4.1. $\square$

Example 4.1 Consider the 3-dimensional table, given in Example 3.2, which is collapsible with respect to $\tau_{12}^{(3)}$. Let $A=\{1,2\}$ so that $A^c=B^c=\{3\}$. It can be checked that

$$\sum_{Z\subseteq\{1,2\}}(-1)^{|A-Z|}\left\{\tilde{l}^{(3)}_{Z\cup\{3\}}(i_Z,i_k)-\tilde{l}^{(3)}_Z(i_Z)\right\}=0$$

and so Eq. (29) is satisfied. Thus, the table is strictly collapsible with respect to $\tau_{12}^{(3)}$.

The following conditions for strict collapsibility are the analogues of Theorems 3.2 and 3.3.

Theorem 4.2 *Let $B \subset \bar{n}$ with $|B| = s < n$, and $A \subseteq B$. An n-dimensional table is strictly collapsible (over B^c) into an s-dimensional table with respect to the set $C_k = \{\tau_L | \{k\} \subseteq L \subseteq A\}$ of interaction factors if and only if*

$$\tilde{d}_A^{(B)}(i_A) = \tilde{d}_{A_k}^{(B)}(i_{A_k}) \tag{35}$$

and

$$l^{(n)}(i) - \tilde{l}_{\bar{n}_k}^{(n)}(i_{\bar{n}_k}) = \tilde{l}_{\bar{s}}^{(n)}(i_{\bar{s}}) - \tilde{l}_{\bar{s}_k}^{(n)}(i_{\bar{s}_k}), \tag{36}$$

where, for example, $\bar{s}_k = \bar{s} \setminus \{k\}$.

Proof The collapsibility follows from Eq. (35) by Theorem 3.2. Let now $C_L = \{Z \mid Z \supseteq L \text{ and } Z \cap B^c \neq \phi\}$. Note that, if $L_1 \subset L_2$, then $C_{L_1} \subseteq C_{L_2}$. Hence, it suffices to show the equivalence of Eq. (36) and the strict collapsibility condition (ii) for $L = \{k\}$. By Theorem 4.1, with $L = \{k\}$,

$$\tau_Z^{(n)}(i_Z) = 0, \quad \text{for all } Z \supseteq \{k\}, Z \cap B^c \neq \phi$$

if and only if

$$\sum_{Z \subseteq L} (-1)^{|L-Z|} \tilde{l}_{Z \cup L^c}^{(n)}(i_Z, i_{L^c}) = \sum_{Z \subseteq L} (-1)^{|L-Z|} \tilde{l}_{Z \cup L^c \setminus B^c}^{(n)}(i_Z, i_{L^c \setminus B^c}). \tag{37}$$

Also, when $L = \{k\}$, $L^c = \bar{n} \setminus \{k\}$ and $L^c \setminus B^c = \bar{s} \setminus \{k\}$. Therefore, Eq. (37) is now equivalent to

$$l^{(n)}(i) - \tilde{l}_{\bar{n}_k}^{(n)}(i_{\bar{n}_k}) = \tilde{l}_{\bar{s}}^{(n)}(i_{\bar{s}}) - \tilde{l}_{\bar{s}_k}^{(n)}(i_{\bar{s}_k}),$$

which completes the proof. □

Example 4.2 Consider the following 3-dimensional table.

X_3	1		2	
	X_2			
X_1	1	2	1	2
1	31.5	141.175	21.115	19.106
2	1.733	1.162	31.5	4.263

It can be shown that

$$\tilde{d}_{12}(i, j) = \begin{cases} 0.713, & \text{if } j = 1, \quad i \in \{1, 2\} \\ 1.127, & \text{if } j = 2, \quad i \in \{1, 2\}; \end{cases} \qquad \tilde{d}_2(j) = \begin{cases} 0.713, & \text{if } j = 1 \\ 1.127, & \text{if } j = 2 \end{cases}$$

which implies

$$\tilde{d}_{12}(i, j) = \tilde{d}_2(j).$$

Hence, the table is collapsible with respect to $\tau_{12}^{(3)}$ and $\tau_1^{(3)}$. Also it can be seen that

$$l^{(3)}(i, j, k) - \tilde{l}_{23}^{(3)}(j, k) = \tilde{l}_{12}^{(3)}(i, j) - \tilde{l}_{2}^{(3)}(j).$$

Hence the table is strictly collapsible into 2-dimensional table with respect to $\tau_1^{(3)}$ and $\tau_{12}^{(3)}$.

Remark 4.1 The necessary and sufficient conditions for strict collapsibility (over B^c) with respect to the set $\{\tau_L | \{j, k\} \subseteq L \subseteq A \subseteq B\}$ can be obtained in the similar way. Indeed, the conditions are

$$\tilde{d}_A^{(B)}(i_A) - \tilde{d}_{A_k}^{(B)}(i_{A_k}) = \tilde{d}_{A_j}^{(B)}(i_{A_j}) - \tilde{d}_{A_{jk}}^{(B)}(i_{A_{jk}}),$$

and

$$\sum_{Z \subseteq \{j,k\}} (-1)^{|Z|} \left\{ \tilde{l}_{\bar{n}\setminus\{j,k\}\cup Z}^{(n)}(i_{\bar{n}\setminus\{j,k\}}, i_Z) - \tilde{l}_{\bar{s}\setminus\{j,k\}\cup Z}^{(n)}(i_{\bar{s}\setminus\{j,k\}}, i_Z) \right\} = 0.$$

4.1 Conditional independence

In this section, we explore the relationship between strict collapsibility and conditional independence. Suppose $X_1, \ldots, X_n$ are n categorical variables. Then X_j and X_k are said to be conditionally independent given the remaining variables if

$$p(i_1, \ldots, i_n) = \frac{\sum_{i_j} p(i_1, \ldots, i_n) \sum_{i_k} p(i_1, \ldots, i_n)}{\sum_{i_j, i_k} p(i_1, \ldots, i_n)}.$$

For an n-dimensional table, it is easy to see that variables X_j and X_k are conditionally independent, given all other variables, if and only if

$$l^{(n)}(i) - \tilde{l}_{\bar{n}_k}^{(n)}(i_{\bar{n}_k}) = \tilde{l}_{\bar{n}_j}^{(n)}(i_{\bar{n}_j}) - \tilde{l}_{\bar{n}_{jk}}^{(n)}(i_{\bar{n}_{jk}}),$$

where $\bar{n}_j = \bar{n} \setminus \{j\}$. A similar result is true for the following more general case. Let $A \cup B \cup C = \bar{n}$ be a partition of $\bar{n}$. Then $X_A \perp X_B \mid X_C$ if and only if

$$l^{(n)}(i) - \tilde{l}_{\bar{n}_A}^{(n)}(i_{\bar{n}_A}) = \tilde{l}_{\bar{n}_B}^{(n)}(i_{\bar{n}_B}) - \tilde{l}_{\bar{n}_{A\cup B}}^{(n)}(i_{\bar{n}_{A\cup B}}). \quad (38)$$

Indeed, one can show that, using the above results, for an n-dimensional table, X_j and X_k are independent, given all other variables, if and only if

$$\sum_{Z \supseteq \{j,k\}} \tau_Z^{(n)}(i_Z) = 0.$$

This result is due to Teugels and Horebeek (1998), which is an extension of a result for conditional independence for the 3-dimensional case due to Birch (1963).

Example 4.3 Consider the following 3-dimensional table:

X_3	1		2		3	
	X_2					
X_1	1	2	1	2	1	2
1	4	2	2	1	1	4
2	2	1	4	2	1	4

For the above table, it can be seen that

$$l_{123}^{(3)}(i,j,k) - \tilde{l}_{13}^{(3)}(i,k) = \tilde{l}_{23}^{(3)}(j,k) - \tilde{l}_{3}^{(3)}(k).$$

Hence, $X_1 \perp X_2 \mid X_3$.

Let $\bar{n} = A \cup B \cup C$, where A,B and C are mutually exclusive. For a hierarchical log-linear model, the n-dimensional table is collapsible into s-dimensional table (over C) with respect to $\tau_{A\cup V}^{(n)}$, where $V \subseteq B$, if $\tau_Z^{(n)} = 0$ for all $Z \cap A \neq \phi$ and $Z \cap C \neq \phi$, that is, if $X_A \perp X_C \mid X_B$.

Bishop et al. (1975, p. 47) stated that the above conditions are necessary and sufficient. Later, Whittemore (1978) showed that they are only sufficient but not necessary.

We next show that those conditions are necessary and sufficient for strict collapsibility with respect to a set of interaction parameters.

Theorem 4.3 *Let $\bar{n} = A \cup B \cup C$ be a partition of $\bar{n}$ such that $|A \cup B| = s$ and $|C| = n - s$. Then, an n-dimensional table is strictly collapsible (over C) into an s-dimensional table with respect to the set $C_L = \{\tau_L | L \subseteq A \cup B; L \cap A \neq \phi\}$ if and only if $X_A \perp X_C \mid X_B$.*

Proof Note that

$$X_A \perp X_C \mid X_B \Longleftrightarrow \tau_Z{=}0, \quad \text{for every Z such that } Z\cap A \neq \phi;\ Z \cap C \neq \phi. \tag{39}$$

Let $k \in A$. We now show that the table is collapsible over $A \cup B$ with respect to τ_L, where $L \subseteq A \cup B$ and $k \in L$. Writing Eq. (38) in terms of interaction factors,

$$l^{(n)}(i) = \sum_{Z\subseteq A\cup B} \tau_Z(i_Z) + \sum_{Z\subseteq B\cup C} \tau_Z(i_Z) - \sum_{Z\subseteq B} \tau_Z(i_Z) \tag{40}$$

which implies

$$p_{A\cup B}(i_{A\cup B}) = \exp\left\{\sum_{Z\subseteq A\cup B} \tau_Z(i_Z) - \sum_{Z\subseteq B} \tau_Z(i_Z)\right\} \sum_{i_j:j\in C} \exp\left\{\sum_{Z\subseteq B\cup C} \tau_Z(i_Z)\right\}.$$

This is equivalent to

$$l^{(s)}(i_{A\cup B}) = \sum_{Z\subseteq A\cup B} \tau_Z(i_Z) - \sum_{Z\subseteq B} \tau_Z(i_Z) + \ln\left(\sum_{i_j:j\in C} \exp\left\{\sum_{Z\subseteq B\cup C} \tau_Z(i_Z)\right\}\right). \tag{41}$$

From Eqs. (40) and (41), we get

$$\tilde{l}^{(n)}_{A\cup B}(i_{A\cup B}) - \tilde{l}^{(n)}_{A_k\cup B}(i_{A_k\cup B}) = l^{(s)}_{A\cup B}(i_{A\cup B}) - \tilde{l}^{(s)}_{A_k\cup B}(i_{A_k\cup B}),$$

where $A_k = A \setminus \{k\}$, and so Eqs. (15) is satisfied. By Theorem 3.2, the table is collapsible with respect to $C_k = \{\tau_L | \{k\} \subseteq L \subseteq A \cup B\}$. Since k is arbitrary, the table is indeed collapsible with respect to τ_L, where $L \subseteq A \cup B$ and $L \cap A \neq \phi$. The strict collapsibility follows again from Eq. (39).

Assume now the table is strictly collapsible so that Eq. (39) holds, which is equivalent to $X_A \perp X_C \mid X_B$. This proves the theorem. $\square$

Corollary 4.1 *Let $k \in \{1, 2\}$. Then a 3-dimensional table is strictly collapsible into a 2-dimensional table with respect to $\tau^{(3)}_k$ and $\tau^{(3)}_{12}$ if and only if $\tau^{(3)}_{123} = 0$ and $\tau^{(3)}_{k3} = 0$.*

Note, for example, when $k = 1$, the conditions $\tau^{(3)}_{123} = 0$ and $\tau^{(3)}_{13} = 0$ are nothing but Bishop et al. (1975) sufficient conditions for collapsibility with respect to $\tau^{(3)}_{12}$ or $\tau^{(3)}_{23}$.

Acknowledgements The authors are grateful to the associate editor and a referee for their helpful and constructive comments, especially for pointing out the use of Möbius inversion for analyzing graphical models. The second author wishes to thank the Council of Scientific and Industrial Research (CSIR), India, for the award of a research fellowship.

References

Andersen, E. B. (1990). *The statistical analysis of categorical data*. Berlin: Springer.

Birch, M. W. (1963). Maximum likelihood in three-way contingency tables. *Journal of the Royal Statistical Society B, 25*, 220–233.

Bishop, Y. M. M., Fienberg, S. E., Holland, P. W. (1975). *Discrete multivariate analysis: theory and practice*. Cambridge: MIT Press.

Charalambides, A. C. (2002). *Enumerative combinatorics*. Florida: Chapman and Hall.

Cox, D. R., Wermuth, N. (2003). A general condition for avoiding effect reversal after marginalization, *Journal of the Royal Statistical Society B, 65*, 937–941.

Ducharme, G. R., Lepage, Y. (1986). Testing collapsibility in multidimensional tables. *Journal of the Royal Statistical Society B, 48*, 197–205.

Guo, J. H., Geng, Z. (1995). Collapsibility of logistic regression coefficient. *Journal of the Royal Statistical Society B, 57*, 263–267.

Hammersley, J. M., Clifford, P. E. (1971). Markov fields on finite graphs and lattices. Unpublished manuscript.

Lauritzen, S. L. (1996). *Graphical models*. New York: Oxford University Press.

Lindley, D. V., Novick, M. R. (1981). The role of exchangeability in inference. *Annals of Statistics, 9*, 45–58.

Shapiro, S. H. (1982). Collapsing contingency tables: a geometrical approach. *American Statistician, 36*, 43–46.

Simonoff, E. H. (2003). *Analyzing categorical data*. New York: Springer.

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

Teugels, J. L., Horebeek, J. V. (1998). Generalized linear models for discrete data. *Statistics & Probability Letters, 38*, 41–47.

Wermuth, N. (1987). Parametric collapsibility and the lack of moderating effects in contingency tables with a dichotomous response variable. *Journal of the Royal Statistical Society B, 49*, 353–364.

Whittaker, J. (1990). *Graphical models in applied multivariate statistics*. Chichester: Wiley.

Whittemore, A. S. (1978). Collapsibility of multidimensional contingency tables. *Journal of the Royal Statistical Society B, 40*, 328–340.