J. R. Statist. Soc. B (1987)
49, *No.* 3, *pp.* 353–364

Parametric Collapsibility and the Lack of Moderating Effects in Contingency Tables with a Dichotomous Response Variable

By NANNY WERMUTH†

University of Mainz, Federal Republic of Germany

[Received May 1986. Final revision May 1987]

SUMMARY

We consider contingency tables having one variable specified as a response with just two categories. We look at conditions for collapsibility of a symmetric and a directed measure of association, the odds-ratio and the relative risk: situations are discussed under which equal partial associations coincide with the corresponding marginal association. Contrary to the odds-ratio the relative risk is collapsible, if there are independencies in the marginal distribution of the influencing variables. This fact is exploited to derive conditions for the lack of a moderating effect, the latter being a much discussed concept in the social science literature.

Keywords: INDEPENDENT REGRESSORS; ODDS-RATIO; RELATIVE RISK; YULE-SIMPSON PARADOX

1. INTRODUCTION

In many empirical studies, one wants to investigate how a response variable with just two categories depends on several qualitative variables. Data from such studies can be summarised in contingency tables containing at least one dichotomous variable, the response. Typically, there are some variables for which potential influences are of main interest and some others that play the role of only *background variables*: one needs to include the latter explicitly in an analysis, because they might modify one of the associations of interest, but one hopes to be able to report the results irrespective of the categories of the background variables. Typical examples are simple multicentre clinical trials, in which the categories of the response variable (V_1) are success or failure of a treatment, those of the background variable (V_3) are the different clinics or sites and those of the main variable of interest (V_2) are the treatment types. Reporting results in such studies irrespective of site can be justified if one has strongly consistent results within sites and these agree with the result overall, that is, for the data combined from all sites. This is a description by example, in which the so-called moderating effect of a background variable is lacking. The concept of moderator variables is much discussed in the social science literature. However, its implementation has remained unsatisfactory. For a survey see Zedeck (1971).

Moderator variables are, in the social sciences, mainly of interest in non-experimental situations like the following. The risk that a person becomes depressive is known to increase with the occurrence of life events like birth, death, severe illness or separation of friends. Since it is, in general, not possible to control such events, it becomes important to understand conditions which are likely to change or moderate this known relationship. Thus, one wants to determine which of the personality characteristics, environmental conditions, coping strategies or therapies are important moderator variables. One wants to estimate moderating effects and to understand under which conditions such effects cannot occur.

Associated statistical tasks are to assess the relative importance of an additional influencing or regressor variable, to estimate the bias introduced by neglecting an important influencing variable and to state conditions under which a measure of association is unaffected by adding

† *Address for correspondence*: Professor N. Wermuth, Psychologisches Institut, Johannes Guttenberg-Universität, Postfach 3980, D-6500 Mainz, Federal Republic of Germany.

 0035-9246/87/49353 $2.00

a variable to an analysis. All of these aspects are well understood in the case of exclusively linear dependencies (Goldberger, 1964, chapter 10). For symmetric associations among discrete variables the last aspect has been called (parametric) *collapsibility* and was studied by Yule (1900), Simpson (1951), Bishop (1971), Whittemore (1978) and Ekholm (1985). In this paper we treat the situation in which all variables under study are discrete and there is a dichotomous response variable.

There are two basic aspects to the *lack of a moderating effect*: (1) associations have to coincide for all levels of the moderator variable and (2) the conclusions regarding the associations have to remain unchanged if the moderator variable is excluded from the analysis. Correspondingly, we speak of *strongly consistent results* in terms of a given measure of association, if all partial associations within sites are equal, and we say that a measure of association is *collapsible* if also the equal partial associations coincide with the marginal association. We discuss in section 2 necessary and sufficient conditions for the collapsibility of the odds-ratio and the relative risk in the simplest case of a 2^3-contingency table. In section 3 we give necessary and sufficient conditions for the lack of a moderating effect in those 2^3-tables, in which equal partial relative risks are always collapsible. Nontrivial sufficient conditions for the lack of a moderating effect on the relative risk are derived in section 4 for general tables with nonzero cell probabilities. Finally, we show in section 5 how, for data in a $2 \times 3 \times 3 \times 2$ table, where the response depends on all three of the remaining variables, the hypothesis of no moderating effect on the relative risk fits the data well.

2. CONDITIONS FOR COLLAPSIBILITY OF ODDS-RATIOS AND OF RELATIVE RISKS

For a $2 \times J \times K$ table of probabilities we use the indices i, j, and k for the categories of the variables V_1, V_2, and V_3, respectively: denote by Π_{ijk} the joint probability $P(V_1 = i, V_2 = j, V_3 = k) > 0$. To make the interpretations of different measures of association more vivid, we name the variables as in simple multicentre studies: V_1 as outcome or response, V_2 as treatment and V_3 as clinic or site. For our discussion on collapsibility we need *marginal probabilities*, such as

$$\Pi_{ij+} = \Sigma_k \Pi_{ijk} = P(V_1 = i, V_2 = j),$$

which are called joint marginal probabilities in the $2 \times J$ *marginal contingency table*. Furthermore, we use *conditional probabilities* such as, for $V_1 = 1$ given $V_2 = j$ and $V_3 = k$,

$$\Pi_{1|jk} = \Pi_{1jk} / \Pi_{+jk}.$$

In the context of log-linear models for contingency tables (Birch, 1963) the natural measures of association are logarithms of so-called odds-ratios. For instance, the odds-ratios for outcome and treatments 1 and 2 within sites are

$$\frac{\Pi_{11k}\Pi_{22k}}{\Pi_{12k}\Pi_{21k}}, \quad k = 1, \ldots, K. \tag{2.1}$$

They are also called *partial odds-ratios* for V_1 and V_2 given V_3.

In epidemiological studies, the measure of association used before the advent of log-linear models was the relative risk (see, e.g., Hill, 1962). For instance, the relative risks for $V_1 = 1$ under treatment 1 compared to treatment 2 within sites are the following ratios of conditional probabilities:

$$\frac{\Pi_{1|1k}}{\Pi_{1|2k}}, \quad k = 1, \ldots, K. \tag{2.2}$$

They are also labelled the *partial relative risks* for $V_1 = 1$, given V_3.

One obtains the *corresponding marginal measures of association* from the probabilities in the marginal table of V_1 and V_2, or, to put it differently, after replacing "k" in (2.1) and (2.2) by

"+". *Results* are said to be *weakly consistent* within sites if the partial associations given sites lead to similar interpretations regarding the success ($V_1 = 1$) of treatment 1 ($V_2 = 1$) in comparison to the success of treatment 2: treatment 1 appears to be better, within sites, than treatment 2 if the odds-ratios in (2.1) or the relative risks in (2.2) are larger than one. Situations in which weakly consistent results appear reversed overall have been named the Yule-Simpson paradox (Yule 1900; Simpson 1951).

If, given treatment 1, response is independent of sites then we speak of *conditional independence* of V_1 *and* V_3 *given* $V_2 = 1$ and write $V_1 \perp\!\!\!\perp V_3 \mid (V_2 = 1)$. In that case one of the following equivalent conditions, stated here without proof, holds:

$$\text{(i)} \qquad \Pi_{1|11} = \Pi_{1|1+}$$

$$\text{(ii)} \qquad \frac{\Pi_{111}\Pi_{212}}{\Pi_{112}\Pi_{211}} = 1 \tag{2.3}$$

$$\text{(iii)} \qquad \frac{\Pi_{1|11}}{\Pi_{1|12}} = 1.$$

For $V_1 \perp\!\!\!\perp V_3 \mid (V_2 = 1)$ *and* $V_1 \perp\!\!\!\perp V_3 \mid (V_2 = 2)$ we write, in brief, $V_1 \perp\!\!\!\perp V_3 \mid V_2$ and speak of *conditional independence* of V_1 *and* V_3 *given* V_2.

The more standard description for $V_1 \perp\!\!\!\perp V_3 \mid V_2$ is $\Pi_{ijk} = \Pi_{ij+}\Pi_{+jk}/\Pi_{+j+}$, and, for $V_3 \perp\!\!\!\perp (V_1, V_2)$, the *complete independence of* V_3 *from* V_2 *and* V_1, taken jointly, it is $\Pi_{ijk} = \Pi_{ij+}\Pi_{++k}$, where both hold for all (i, j, k). It is known (see, e.g., Birch, 1963) that $V_3 \perp\!\!\!\perp (V_1, V_2)$ is equivalent to $V_1 \perp\!\!\!\perp V_3 \mid V_2$ and $V_2 \perp\!\!\!\perp V_3 \mid V_1$. We now turn to conditions for collapsibility.

Proposition 1. In 2^3-contingency tables necessary and sufficient conditions for results, which are strongly consistent, also to be collapsible are

(i) for odds-ratios of V_1 and V_2 (2.1): $V_1 \perp\!\!\!\perp V_3 \mid V_2$ or $V_2 \perp\!\!\!\perp V_3 \mid V_1$,

(ii) for relative risks of V_1 (2.2): $V_1 \perp\!\!\!\perp V_3 \mid V_2$ or $V_2 \perp\!\!\!\perp V_3$.

Statement (i) is an immediate consequence of the conditions for collapsibility of log-linear interaction parmeters (Whittemore, 1978), and (ii) is proven in the Appendix.

Remark 1. As far as collapsibility is concerned, odds-ratios behave analogously to concentrations σ^{ij}. These are elements in the inverse of a positive definite covariance matrix Σ and they are simple multiples of partial correlation coefficients given all other variables (see, e.g., Wermuth, 1976). For trivariate normal distributions the following three statements are equivalent:

(i) $V_i \perp\!\!\!\perp V_j \mid V_k$,

(ii) $\sigma^{ij} = 0$,

(iii) $\rho_{ij.k} = 0$,

where i, j and k are distinct elements of $\{1, 2, 3\}$ and $\rho_{ij.k}$ denotes a partial correlation coefficient. The marginal concentration between V_1 and V_2, $\sigma^{12.3}$, is known to be related to the partial one, σ^{12} (see, e.g., Wermuth, 1980) through

$$\sigma^{12.3} = \sigma^{12} - \sigma^{13}\sigma^{23}/\sigma^{33},$$

so that, from $\sigma^{33} > 0$, by positive definiteness, the claimed analogy becomes evident. This is not surprising, if one considers the exponential family parametrisations (see, e.g., Dempster, 1971) of a joint normal and of a multinomial distribution for three variables. Concentrations are there the natural parameters for the former and log-odds ratios for the latter, provided the prerequisite for collapsibility is satisfied, that is, provided one has equal partial odds-ratios.

Remark 2. Relative risks behave similarly, but not completely analogously, to regression coefficients in bivariate normal regressions. Let $\beta_{12.3}$ denote the partial regression coefficient, the coefficient of V_2 in a linear regression of V_1 on V_2 and V_3, and β_{12}, the marginal regression coefficient, the coefficient of V_2 in a linear regression of V_1 on V_2 alone. Then it is known (Currie and Korabinski, 1984) that the two are related through

$$\beta_{12.3} = \beta_{12} \frac{(\rho_{12} - \rho_{13}\rho_{23})}{\rho_{12}(1 - \rho_{23}^2)},$$

where the ρ's are marginal correlation coefficients. This makes it plain that the regression coefficient is collapsible in the case of $\beta_{12.3} \neq 0$ (not $V_1 \perp\!\!\!\perp V_2 \mid V_3$) if and only if

$$\rho_{23} = 0 \quad \text{or} \quad \rho_{13} = \rho_{12}\rho_{23}$$

which for joint normal distributions is equivalent to requiring that

$$V_2 \perp\!\!\!\perp V_3 \quad \text{or} \quad V_1 \perp\!\!\!\perp V_3 \mid V_2.$$

The difference, which remains to relative risks, concerns the strong consistency in terms of the analogous measures: concentrations and odds-ratios. In normal distribution theory the conditional associations between V_1 and V_2, given different levels of V_3, are equal by definition, which is reflected in just one possible value for the partial concentration σ^{12} (or the partial correlation coefficient $\rho_{12.3}$). The relative risk can however, for $V_2 \perp\!\!\!\perp V_3$, be collapsible even if one has only weakly consistent results in terms of odds-ratios. An example is given in Table 1.

TABLE 1

Example for collapsible relative risks in spite of unequal partial odds-ratios*

	$k=1$		$k=2$		*Sum over k*	
	$j=1$	$j=2$	$j=1$	$j=2$	$j=1$	$j=2$
$i=1$	6 (30%)	40 (20%)	30 (75%)	200 (50%)	36 (60%)	240 (40%)
$i=2$	14	160	10	200	24	360
Sum	20	200	40	400	60	600
	$\frac{\Pi_{1\|11}}{\Pi_{1\|21}} = 1.5$		$\frac{\Pi_{1\|12}}{\Pi_{1\|22}} = 1.5$		$\frac{\Pi_{1\|1+}}{\Pi_{1\|2+}} = 1.5$	

* All probabilities Π_{ijk} are multiplied by 660.

Remark 3. If outcome is independent of treatment, given sites ($V_1 \perp\!\!\!\perp V_2 \mid V_3$), then the results are strongly consistent in terms of relative risks and in terms of odds-ratios, and either both or none of these measures of associations are collapsible. Examples for the latter case are special versions of the Yule-Simpson paradox described for instance by Birch (1963, (5.1)) or by Bishop, Fienberg and Holland (1975, p. 41). To see the former, note that $V_1 \perp\!\!\!\perp V_2 \mid V_3$ *and* $V_2 \perp\!\!\!\perp V_3 \mid V_1$ imply $V_2 \perp\!\!\!\perp (V_3, V_1)$ and hence $V_2 \perp\!\!\!\perp V_3$. Similarly, $V_1 \perp\!\!\!\perp V_2 \mid V_3$ and $V_2 \perp\!\!\!\perp V_3$ imply $V_2 \perp\!\!\!\perp (V_1, V_3)$, and in particular $V_2 \perp\!\!\!\perp V_3 \mid V_1$.

Remark 4. If outcome depends conditionally on both of treatments and sites, then at most

one of the measures of association is collapsible. To see this, note first that, for $V_1 \perp\!\!\!\perp V_2 \mid V_3$ not to be satisfied, strongly consistent results in terms of odds-ratios and relative risks imply $V_1 \perp\!\!\!\perp V_3 \mid V_2$ (for details see e.g., Wermuth, 1986) and hence contradict an assumption. Note secondly that the two conditions for collapsibility of odds-ratios and relative risks cannot both be satisfied in tables with neither $V_1 \perp\!\!\!\perp V_2 \mid V_3$ nor $V_1 \perp\!\!\!\perp V_3 \mid V_2$: Birch (1963) has shown that in $2 \times J \times K$ tables the two independencies $V_2 \perp\!\!\!\perp V_3 \mid V_1$ and $V_2 \perp\!\!\!\perp V_3$ together imply that one of the stronger hypotheses $V_2 \perp\!\!\!\perp (V_1, V_3)$ or $V_3 \perp\!\!\!\perp (V_1, V_2)$ has to hold. The claim then follows, because $V_1 \perp\!\!\!\perp V_2 \mid V_3$ and $V_1 \perp\!\!\!\perp V_3 \mid V_2$ are necessary conditions for $V_2 \perp\!\!\!\perp (V_1, V_3)$ and for $V_3 \perp\!\!\!\perp (V_1, V_2)$, respectively.

One conclusion, for data analyses, from Proposition 1 and the example in Table 1 is that collapsibility of relative risks offers new possibilities for giving more condensed descriptions of results in situations where the partial odds-ratios do not coincide. We therefore investigate tables satisfying $V_2 \perp\!\!\!\perp V_3$ next.

3. EFFECTS OF ONE MARGINAL INDEPENDENCE IN A 2^3-TABLE

A 2^3-table with $V_2 \perp\!\!\!\perp V_3$ can be viewed as the simplest situation in which a dichotomous response variable (V_1) has two qualitative variables as possible influences and these are independent of each other. For $V_2 \perp\!\!\!\perp V_3$ one has, by definition, for $j = 1, 2$ and $k = 1, 2$,

(i) $$\Pi_{+jk}/\Pi_{+j+} = \Pi_{++k},$$

or

(ii) $$\frac{\Pi_{+11}\Pi_{+22}}{\Pi_{+21}\Pi_{+12}} = 1, \tag{3.1}$$

or

(iii) $$\Pi_{+1k}/\Pi_{+1+} = \Pi_{+2k}\Pi_{+2+}.$$

This restriction on the probabilities in the marginal contingency table of V_2 and V_3 has effects on the possible associations in the joint table, to see which, we note first that then

$$\frac{\Pi_{1|1+}}{\Pi_{1|2+}} = a\,\frac{\Pi_{1|11}}{\Pi_{1|21}} + (1-a)\,\frac{\Pi_{1|12}}{\Pi_{1|22}}, \tag{3.2}$$

with $a = \Pi_{1|21}\Pi_{++1}/(\Pi_{1|21}\Pi_{++1} + \Pi_{1|22}\Pi_{++2})$.

Thus, contrary to the case for general 2^3-tables, the marginal relative risk is a weighted average of the corresponding partial quantities, with positive weights adding to one. The implications of (3.2) are (1) that strongly consistent results in terms of relative risks are always collapsible, i.e., in agreement with the result overall, and (2) that weakly consistent results within sites can never appear reversed overall, that is, the Yule-Simpson paradox cannot occur. The last statement follows after noting that, in tables with a dichotomous response, weak consistency in terms of one measure of association implies weak consistency in terms of the other.

Proposition 2: In 2^3-contingency tables with $V_2 \perp\!\!\!\perp V_3$ the necessary and sufficient condition for the lack of a moderating effect of V_3 is

(i) for odds-ratios of V_1 and V_2: $V_2 \perp\!\!\!\perp V_3 \mid V_1$,

(ii) for relative risks of $V_1 = 1$: $V_2 \perp\!\!\!\perp V_3 \mid (V_1 = 1)$.

A moderating effect is lacking, if one has equal partial associations and these are collapsible. Since $V_2 \perp\!\!\!\perp V_3$ is sufficient for collapsibility of equal relative risks, Proposition 2 (ii) results by

noting the following equivalent statements for 2^3-tables with $V_2 \perp\!\!\!\perp V_3$:

$$\text{(i)} \quad \frac{\Pi_{1|11}}{\Pi_{1|21}} = \frac{\Pi_{1|12}}{\Pi_{1|22}},$$

$$\text{(ii)} \quad \frac{\Pi_{111}\Pi_{122}}{\Pi_{121}\Pi_{112}} = 1, \tag{3.3}$$

$$\text{(iii)} \quad V_2 \perp\!\!\!\perp V_3 \,|\, (V_1 = 1).$$

For the proof of Proposition 2 (i) note first from Proposition 1 that conditions for the lack of a moderating effect on the odds-ratio are included in those for the collapsibility: if $V_1 \perp\!\!\!\perp V_3 \,|\, V_2$ or $V_2 \perp\!\!\!\perp V_3 \,|\, V_1$ then one has equal partial odds-ratios. Note secondly that these conditions are still necessary and sufficient in 2^3-tables with $V_2 \perp\!\!\!\perp V_3$, but that conditional independencies of variable pairs have different interpretations,

$$\text{(i)} \quad V_1 \perp\!\!\!\perp V_2 \,|\, V_3 \text{ if and only if } V_2 \perp\!\!\!\perp (V_1, V_3),$$

$$\text{(ii)} \quad V_1 \perp\!\!\!\perp V_3 \,|\, V_2 \text{ if and only if } V_3 \perp\!\!\!\perp (V_1, V_2), \tag{3.4}$$

$$\text{(iii)} \quad V_2 \perp\!\!\!\perp V_3 \,|\, V_1 \text{ if and only if } V_2 \perp\!\!\!\perp (V_1, V_3) \text{ or } V_3 \perp\!\!\!\perp (V_1, V_2).$$

The only difficult part of (3.4) is, in (iii), the implication of $V_2 \perp\!\!\!\perp V_3$ and $V_2 \perp\!\!\!\perp V_3 \,|\, V_1$. It follows, however, from one of Birch's results (1963, (5.3)) and it completes the proof of Proposition 2.

Remark 1: The hypotheses $V_2 \perp\!\!\!\perp V_3 \,|\, (V_1 = 1)$ and $V_2 \perp\!\!\!\perp V_3 \,|\, V_1$, which are of little interest in general 2^3-tables with V_1 as a response variable, become important in 2^3-tables having marginal independence of treatment (V_2) and site (V_3).

Remark 2: The hypothesis of strongly consistent results in terms of relative risks may be tested with standard methods for odds-ratios (Bishop, Fienberg and Holland, 1975, p. 494).

Remark 3: Equal partial odds-ratios are never collapsible in 2^3-tables with $V_2 \perp\!\!\!\perp V_3$ and with outcome (V_1) being marginally dependent on both of treatment (V_2) and site (V_3). This follows from (3.4) (iii) and Proposition 1. The further equivalencies in (3.4) (i) and (ii) imply, together with Proposition 2, that V_3 has a moderating effect on the odds-ratio in all situations except in the more clear-cut ones, in which either outcome is independent of treatments given sites, (3.4) (i), or outcome is independent of sites given treatments, (3.4) (ii).

Remark 4: If probabilities are estimated for 2^3-tables satisfying $V_2 \perp\!\!\!\perp V_3$, then tables with $V_2 \perp\!\!\!\perp V_3 \,|\, (V_1 = i)$ for fixed, i are likely to show a reasonable fit to the hypothesis $V_2 \perp\!\!\!\perp V_3 \,|\, V_1$ as well. One can then use the equivalence in (3.4) (iii) to discriminate between the two hypotheses: if statistical tests show V_1 to depend conditionally or marginally on both of V_2 and V_3, then $V_2 \perp\!\!\!\perp V_3 \,|\, V_1$ should be rejected in favour of the less restrictive hypothesis $V_2 \perp\!\!\!\perp V_3 \,|\, (V_1 = i)$ for a fixed i.

Proposition 3: In 2^3-contingency tables with outcome (V_1) depending on sites (V_3) given treatments (V_2) the hypothesis of no moderating effect of site on the relative risk for $V_1 = 1$ is equivalent to $V_2 \perp\!\!\!\perp V_3$ and $V_2 \perp\!\!\!\perp V_3 \,|\, (V_1 = 1)$. This result is an immediate consequence of Propositions 1 and 2.

Remark. The combination of the two hypotheses in Proposition 3 does not fit into the log-linear model framework, but it suggests the generalisations treated in the next section.

4. THE LACK OF A MODERATING EFFECT ON RELATIVE RISKS

The results of the previous two sections can be exploited to derive nontrivial sufficient conditions for the lack of a moderating effect of one background variable V_M on the relative risk in general contingency tables having only nonzero cell probabilities Π. Let V_1 be the response, V_2 a treatment variable, V_M be the moderator variable and $V_3' = V_3 \cup \ldots \cup V_{M-1}$ the remaining variables, with I, J, L and K categories, respectively. Then the partial relative risks for $V_1 = 1$ under treatment j ($V_2 = j$) compared to treatment j' ($V_2 = j'$), given V_3' and V_M, are

$$\frac{\Pi_{1|jkl}}{\Pi_{1|j'kl}} = \frac{\Pi_{1jkl}/\Pi_{+jkl}}{\Pi_{1j'kl}/\Pi_{+j'kl}}, \text{ all } k, l \tag{4.1}$$

Proposition 4: In an $I \times J \times K \times L$-contingency table with $\Pi_{ijkl} > 0$, sufficient conditions for the collapsibility over V_M of the partial relative risks (4.1) that are equal for all levels of V_M given $V_3' = k$ are

$$V_2 \perp\!\!\!\perp V_M \mid V_3' \quad \text{or} \quad V_1 \perp\!\!\!\perp V_M \mid (V_3', V_2).$$

This follows from the generalisation of the necessary and sufficient condition (A-2) in a 2^3-table to the general case with (4.1) and $j = 1, j' = 2$:

$$\Sigma_{l \neq l'} (\Pi_{1|2kl} - \Pi_{1|2kl'})(\Pi_{+1kl}\,\Pi_{+2kl'} - \Pi_{+2kl}\Pi_{+1kl'}) = 0$$

and the appropriate generalisations of (2.3) (iii) and (3.1) (ii).

Proposition 5: In an $I \times J \times K \times L$-contingency table with $\Pi_{ijkl} > 0$, V_M has no moderating effect on the relative risk (4.1)

(i) given $V_3' = l$ if $V_2 \perp\!\!\!\perp V_M \mid (V_3', V_1 = 1)$ *and* $V_2 \perp\!\!\!\perp V_M \mid V_3'$,

(ii) irrespective of V_3' if $V_2 \perp\!\!\!\perp (V_M, V_3') \mid (V_1 = 1)$ *and* $V_2 \perp\!\!\!\perp (V_M, V_3')$.

For the proof of Proposition 5 (i), note that the two independencies hold if and only if

$$\Pi_{1jkl} = \Pi_{1jk+}\,\Pi_{1+kl}/\Pi_{1+k+} \quad \text{and} \quad \Pi_{+jkl} = \Pi_{+jk+}\,\Pi_{++kl}/\Pi_{++k+}.$$

These imply that, for $j = 1, j' = 2$,

$$\frac{\Pi_{1|1kl}}{\Pi_{1|2kl}} = \frac{\Pi_{11k+}/\Pi_{+1k+}}{\Pi_{12k+}/\Pi_{+2k+}} = \frac{\Pi_{1|1k+}}{\Pi_{1|2k+}},$$

so that the partial relative risks (4.1) are equal for all levels l of V_M given V_3' and collapsible over V_M.

Similarly, note that the two independencies in 5 (ii) are satisfied if and only if

$$\Pi_{1jkl} = \Pi_{1j++}\,\Pi_{1+kl}/\Pi_{1+++} \quad \text{and} \quad \Pi_{+jkl} = \Pi_{+j++}\,\Pi_{++kl}.$$

These give, for $j = 1, j' = 2$,

$$\frac{\Pi_{1|1kl}}{\Pi_{1|2kl}} = \frac{\Pi_{11++}/\Pi_{+1++}}{\Pi_{12++}/\Pi_{+2++}} = \frac{\Pi_{1|1++}}{\Pi_{1|2++}},$$

so that the partial relative risks (4.1) are equal for all levels of V_3 and V_M and collapsible over V_3' and V_M. This completes the proof.

Lemma: In an $I \times J \times K \times L$-contingency table with $\Pi_{ijkl} > 0$ following a multinomial distribution for a given total number of observations,

(i) $V_2 \perp\!\!\!\perp V_M \mid V_3'$ *and* $V_2 \perp\!\!\!\perp V_M \mid (V_3', V_1)$ if and only if in addition either of $V_1 \perp\!\!\!\perp V_2 \mid (V_3', V_M)$ or $V_1 \perp\!\!\!\perp V_M \mid (V_2, V_3')$ holds

(ii) $V_2 \perp\!\!\!\perp V_M \mid V_3'$ *and* $V_2 \perp\!\!\!\perp V_M \mid (V_3', V_1 = 1)$ (i.e. V_M has moderating effect on the relative risk (4.1) given the levels of V_3') can be true, if no other independencies are satisfied.

The first part (i) is a special case of Proposition 3 in Wermuth and Lauritzen (1983) and (ii) is proven with the estimated probabilities for the data presented in the next section.

5. A SET OF DATA WITH OBSERVED RELATIVE RISKS BEING NEARLY COLLAPSIBLE

From a follow-up investigation by Spielberger on smoking habits of students (Spielberger *et al.*, 1984) observations on 2317 persons are summarised in a $2 \times 3 \times 3 \times 2$-table, where the variables are as follows,

V_1: smoking habits of the respondent

V_2: older sibling as role model

V_3: parents as role model

V_4: sex of the respondent.

The first variable is a dichotomous response variable, V_2 and V_3 are the main influences of interest and V_4 is a background variable.

The first two columns in Table 2 show the observed frequencies n_{1jkl} and $n_{+jkl} = \Sigma_i n_{ijkl}$. The third column contains the observed risks to smoke (i.e. for $V_1 = 1$), $p_{1|jkl} = n_{1jkl}/n_{+jkl}$. Those reflect what is supported by standard log-linear analyses, namely that smoking habits of the students (V_1) depend conditionally on all three of V_2, V_3, V_4; compare Table 3.
From the hypotheses related to sex (V_4) having no moderating effect on the relative risks to

TABLE 2

Observed and estimated risks to smoke for 2317 students*

Sex of respondent	*Smoking behaviour of parents*	*Older siblings*	*Observed frequencies* n_{ijkl}	n_{+jkl}	*Observed risks to smoke* $p_{1\|jkl}$	*Estimated risks to smoke* $\tilde{p}_{1\|jkl}$	*Estimated relative risks* *Siblings' effects* $\frac{\tilde{p}_{1\|j'kl}}{p_{1\|2kl}}$	*Parents' effects* $\frac{\tilde{p}_{1\|jk'l}}{\tilde{p}_{1\|j2l}}$
female	none smokes	no	1	4	.250	.223	1.09	.55
		do not smoke	37	172	.215	.205		.58
		smoke	26	52	.500	.545	2.66	.93
	one smokes	no	8	18	.444	.407	1.16	
		do not smoke	61	178	.343	.351		
		smoke	103	174	.592	.587	1.67	
	both smoke	no	20	34	.588	.466	1.00	1.14
		do not smoke	97	191	.508	.465		1.32
		smoke	222	311	.714	.750	1.61	1.28
male	none smokes	no	1	7	.143	.163	1.09	.46
		do not smoke	36	254	.142	.149		.49
		smoke	21	49	.429	.396	2.66	.77
	one smokes	no	5	16	.313	.354	1.16	
		do not smoke	48	152	.316	.306		
		smoke	83	164	.506	.511	1.67	
	both smoke	no	9	32	.281	.402	1.00	1.14
		do not smoke	61	169	.361	.403		1.32
		smoke	233	340	.685	.650	1.61	1.27

* Estimates as defined in (4.1)

TABLE 3
Selected likelihood-ratio chi-squared test results for the data in Table 2

Variables in the table	*Conditional independence given all remaining variables for pair*	χ^2*-value*	*Degrees of freedom*	*p-value*
(V_1, V_2, V_3, V_4)	(V_1, V_2)	150.0	12	0.000
	(V_1, V_3)	104.5	12	0.000
	(V_1, V_4)	22.9	9	0.006
	(V_2, V_3)	222.6	16	0.000
	(V_2, V_4)	15.7	12	0.205
	(V_3, V_4)	18.8	12	0.094
(V_2, V_3, V_4)	(V_2, V_3)	322.6	8	0.000
	(V_2, V_4)	7.3	6	0.291
	(V_3, V_4)	20.7	6	0.002

smoke (compare Proposition 5), it turns out that only $V_4 \perp\!\!\!\perp V_2 \mid V_3$ in the marginal $3 \times 3 \times 2$ table of (V_2, V_3, V_4) and $V_4 \perp\!\!\!\perp V_2 \mid (V_3, V_1)$ in the joint table of (V_1, V_2, V_3, V_4) fit reasonably though not well. These results and the Lemma, however, reveal that both of the hypotheses cannot be satisfied but that instead $V_4 \perp\!\!\!\perp V_2 \mid V_3$ and $V_4 \perp\!\!\!\perp V_2 \mid (V_3, V_1 = 1)$ may hold, that is, for known parental smoking patterns (V_3), sex of the respondent has no moderating effect on the relative risks obtained by comparing situations that differ with respect to the role model older siblings provide.

We proceed by computing smoothed estimates of risks as follows. First, the maximum-likelihood estimates $\hat{m}_{ijkl}$ of the cell frequencies m_{ijkl} are obtained under the independence hypothesis $V_4 \perp\!\!\!\perp V_2 \mid V_3$ on the (V_2, V_3, V_4)-margin (compare Birch, 1963). These are then modified to satisfy the additional hypothesis $V_4 \perp\!\!\!\perp V_2 \mid (V_3, V_1 = 1)$. More precisely we compute for the data at hand,

$$\text{(i)} \quad \hat{m}_{ijkl} = \frac{n_{ijkl} n_{+jk+} n_{++kl}}{n_{+jkl} n_{++k+}}$$

$$\text{(ii)} \quad \tilde{m}_{1jkl} = \frac{\hat{m}_{1jk+} \hat{m}_{1+kl}}{\hat{m}_{1+k+}}, \quad \tilde{m}_{2jkl} = \hat{m}_{+jkl} - \tilde{m}_{1jkl} \tag{4.1}$$

$$\text{(iii)} \quad \tilde{\Pi}_{1|jkl} = \frac{\tilde{m}_{1jkl}}{\tilde{m}_{+jkl}}.$$

The probability estimates $\tilde{\Pi}_{1|jkl}$ fit both of the hypotheses $V_4 \perp\!\!\!\perp V_2 \mid (V_3, V_1 = 1)$ and $V_4 \perp\!\!\!\perp V_2 \mid V_3$.

The estimates are reasonable in the sense that they provide a good approximation to the data in Table 3 and permit a more condensed description of the effects, since they fit conditions for the lack of a moderating effect of the background variable to the data. The most outstanding result for these data concerns families, in which parents do not smoke $(V_3 = 1)$: if a student has older siblings who smoke $(V_2 = 3)$, then the risk that *he or she* $(V_4 = 1$ or $V_4 = 2)$ will also smoke is estimated to be 166% higher than for a student whose older siblings do not smoke $(V_2 = 2)$. This is reflected in the relative risk of 2.66 and holds even though the data convey a clear conditional dependence of the smoking behaviour (V_1) on the sex of the student (V_4): the risks to smoke are higher for females than for males at all levels of the remaining variables (V_1, V_3). The lack of a moderating effect of V_4 on the relative risk of $V_1 = 1$ for V_2 implies that one can estimate the effects of older siblings more stably than others estimated from the four-way table: they are in fact based on observations in a marginal table, that of V_1, V_2 and V_3. However, the statistical properties of our estimates in (4.1) still need to be investigated

and a formal goodness-of-fit test has to be developed. In a forthcoming paper it will be shown how our estimates relate to conditional (Andersen, 1973) and overall maximum-likelihood estimates.

6. DISCUSSION

During the last twenty years, the statistical literature on contingency tables strongly emphasised the use of log-linear models for analysing structures. Within this framework marginal independencies are, in general, not easily incorporated (see Birch, 1963). As a consequence, tables of the type discussed in this paper have received little or no attention and their nice properties have not been noticed. Model classes for contingency tables that permit one to impose structure on marginal tables, as well, only seem to have been proposed by Goodman (1973), Wermuth and Lauritzen (1983), Kiiveri, Speed and Carlin (1984) and Lauritzen and Wermuth (1984).

Collapsibility in contingency tables has been defined either for probabilities (Lauritzen, 1982; Asmussen and Edwards, 1983) or for interaction parameters (see Simpson, 1951; Darroch, 1963; Plackett, 1969; Bishop, 1971; Whittemore, 1978). The former proved helpful in answering questions about the equivalence of models. However, the Yule-Simpson paradox can occur when conditions for collapsibility in probabilities are satisfied (examples have been given by Wermuth, 1986). This concept was neither designed to study conditions under which parameters measuring associations, other than the probabilities, may correctly be evaluated in a marginal table, nor is it appropriate for this purpose.

We have shown that conditions for collapsibility of relative risks in tables having a dichotomous response variable involve independencies in the marginal distribution of the variables influencing the response (Proposition 1). The opposite pole to collapsibility of associations or interactions in contingency tables is the Yule-Simpson paradox, and this is caused by strong associations among the influences or regressor variables. The Yule-Simpson paradox can be seen as analogous to problems discussed, in other contexts, under the name of multicollinearity in multiple linear regression (Goldberger, 1964) or as the effects of nonorthogonal factors in analyses of variance (Snedecor and Cochran, 1967). Conditions for the absence of the Yule-Simpson paradox coincide with those under which relative risks possess nice properties: in the case of independent regressors, the marginal relative risk is a simple weighted average of the partial relative risks, equal partial relative risks are always collapsible (3.2), and the hypothesis of equal partial relative risks is equivalent to a conditional independence (3.3).

In view of these results it no longer seem necessary to only approximate relative risks by odds-ratios, as has been proposed by Armitage (1975). One can plan studies in such a way that the resulting data fit hypotheses on marginal independencies perfectly. A decision on this has, of course, to depend on the purpose of the study: if in the target population of the study no marginal independencies are expected, then this can be an important argument against forcing them upon any sample. Many studies are, however, for reasons not discussed here, designed in such a way that a main variable of interest and a background variable behave like independent regressor variables: for instance, all multicentre clinical trials, in which equal numbers of patients are assigned to the treatments within sites, and all multicentre retrospective studies having equal numbers of patients with and without a given disease at all sites. Having equal numbers of patients within sites is just a special case of proportional allocation of observable units to treatments (V_2) within sites (V_3), and the latter is the empirical counterpart of the marginal independence of V_2 and V_3. Given such data, one can test the hypothesis that site has no moderating effect on the relative risk with standard methods (compare Proposition 2).

For more general types of data one can still relate the hypotheses of collapsibility over site and of no moderating effect of site on the relative risk to conditional independencies (Propositions 3, 4 and 5), and these permit us to reconcile seemingly contradictory results in data analysis (Compare the Lemma and Tables 2 and 3).

ACKNOWLEDGEMENT

I want to thank Peter McCullagh, Paul Holland, Inge and Hanns-Georg Leimer for helpful comments on an earlier version of this paper and a referee, whose concise remarks helped me to strengthen and generalise the results.

APPENDIX

Here we prove Proposition 1 (ii). We assume again that all probabilities Π_{ijk} are positive.

In a 2^3-table with equal partial relative risks for $V_1 = 1$ given V_3, this relative risk is collapsible over V_3 if and only if $V_1 \perp\!\!\!\perp V_3 \mid V_2$ or $V_2 \perp\!\!\!\perp V_3$.

Note that the marginal risk for $V_1 = 1$ given $V_2 = j$ is, for $j = 1, 2$,

$$\begin{aligned}\Pi_{1|j+} &= \Pi_{1j+}/\Pi_{+j+} \\ &= (\Pi_{1|j1}\Pi_{+j1} + \Pi_{1|j2}\Pi_{+j2})/\Pi_{+j+}. \end{aligned} \tag{A-1}$$

With equal partial relative risks for $V_1 = 1$ given V_3 we have

$$\frac{\Pi_{1|11}}{\Pi_{1|21}} = \frac{\Pi_{1|12}}{\Pi_{1|22}}$$

and can express the marginal relative risk, by using (A-1), as

$$\frac{\Pi_{1|1+}}{\Pi_{1|2+}} = \frac{\Pi_{1|11}}{\Pi_{1|21}} \frac{(\Pi_{1|21}\Pi_{+11} + \Pi_{1|22}\Pi_{+12})\Pi_{+2+}}{(\Pi_{1|21}\Pi_{+21} + \Pi_{1|22}\Pi_{+22})\Pi_{+1+}}$$

This shows that the marginal relative risk equals the partial one if and only if

$$(\Pi_{1|21}\Pi_{+11} + \Pi_{1|22}\Pi_{+12})(\Pi_{+21} + \Pi_{+22}) = (\Pi_{1|21}\Pi_{+21} + \Pi_{1|22}\Pi_{+22})(\Pi_{+11} + \Pi_{+12}).$$

After rewriting we obtain

$$(\Pi_{1|21} - \Pi_{1|22})(\Pi_{+11}\Pi_{+22} - \Pi_{+21}\Pi_{+12}) = 0, \tag{A-2}$$

which is satisfied if and only if

$$\frac{\Pi_{1|21}}{\Pi_{1|22}} = 1 \quad \text{or} \quad \frac{\Pi_{+11}\Pi_{+22}}{\Pi_{+21}\Pi_{+12}} = 1.$$

These two conditions are, from (2.3) (iii) and (3.1) (ii), seen to be equivalent to $V_1 \perp\!\!\!\perp V_3 \mid (V_2 = 2)$ and $V_2 \perp\!\!\!\perp V_3$, respectively. With the assumption of equal partial risks, the first implies $V_1 \perp\!\!\!\perp V_3 \mid V_2$ and the proof is complete.

REFERENCES

Andersen, E. B. (1973) *Conditional Inference and Models for Measuring*. København: Mentalhygiejnisk Forlag.

Armitage, P. (1975) The use of the cross-product ratio in aetiological survey. In: *Perspectives in Probability and Statistics* edited by J. M. Gani), pp. 343–355. London, New York: Academic Press.

Asmussen, S. and Edwards, D. (1983) Collapsibility and response variables in contingency tables. *Biometrika*, **70**, 567–578.

Birch, M. W. (1963) Maximum likelihood in three-way contingency tables. *J. R. Statist. Soc.* B, **25**, 220–233.

Bishop, Y. M. M. (1971) Effects of collapsing multidimensional contingency tables. *Biometrics*, **27**, 545–562.

Bishop, Y. M. M., Fienberg, S. E. and Holland, P. W. (1975) *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, Mass. and London: MIT Press.

Currie, I. and Korabinski, A. (1984) Some comments on bivariate regression. *The Statistician*, **33**, 283–293.

Darroch, J. N. (1962) Interactions in multifactor contingency tables. *J. Royal Statist. Soc. B*, **24**, 251–263.

Dempster, A. P. (1971) An overview of multivariate data analysis. *J.Multivar. Anal.*, **1**, 316–46.

Ekholm, A. (1985) A recursion formula for the log linear parameters of a collapsed contingency table. Research Report 53, Department of Statistics,University of Helsinki.

Golberger, A. S. (1964) *Econometric Theory*. New York: Wiley.

Goodman, L. A. (1973) The analysis of contingency tables, when some variables are posterior to others: a modified path analysis approach. *Biometrika*, **60**, 179–192.
Hill, A. B. (1962) Statistical Methods in Clinical and Preventive Medicine. New York: Oxford Univ. Press.
Kiiveri, H., Speed, T. P. and Carlin, J. B. (1984) Recursive causal models, *J. Australian Math. Soc.* A, **36**, 30–52.
Lauritzen, S. L. (1982) *Lectures on Contingency Tables*, 2nd printing. Aalborg: Aalborg Univ. Press.
Lauritzen, S. L. and Wermuth, N. (1984) Mixed interaction models. Research report, Aalborg Universitets Center (ISSN-0106-0791).
Plackett, R. L. (1969) Multidimensional contingency tables: a survey of models and methods. *Bull. Int. Statist. Inst.*, **43**, 133–141.
Simpson, E. H. (1951) The interpretation of interaction in contingency tables. *J. R. Statist. Soc.* B, **13**, 238–241.
Snedecor, G. W. and Cochran, W. G. (1967) *Statistical Methods*, 6th edition. Ames: Iowa State University Press.
Spielberger, C. D., Jacobs, G. A., Crane, R. S. and Russell, S. F. (1983) On the relation between family smoking habits and the smoking behaviour of college students. *Int. Rev. Appl. Psychol.*, **32**, 54–69.
Wermuth, N. (1976) Analogies between multiplicative models in contingency tables and covariance selection. *Biometrics*, **32**, 95–108.
———(1980) Linear recursive equations, covariance selection, and path analysis. *J. Amer. Statist. Ass.*, **75**, 963–97.
———(1986) Implications of the Yule-Simpson paradox for simple multicentre studies. In: *Research Report 86-1, Stochastik und verwandte Gebiete*, University of Mainz. (ISSN 0177-0098).
Wermuth, N. and Lauritzen, S. L. (1983) Graphical and recursive models for contingency tables, *Biometrika*, **70**, 537–552.
Whittemore, A. S. (1978) Collapsibility of multidimensional contingency tables. *J. R. Statist. Soc.* B, **40**, 328–340.
Yule, G. U. (1900) On the association of attributes in statistics. *Phil. Trans. Ser* A., **194**, 257–319.
Zedeck, S. (1971) Problems with the use of "moderator" variables. *Psychol. Bull.*, **76**, 295–310.