Metrika
https://doi.org/10.1007/s00184-024-00976-y

On the comparison of regression coefficients across multiple logistic models with binary predictors

Luca La Rocca[1] · **Monia Lupparelli**[2] · **Alberto Roverato**[3]

Received: 27 December 2023 / Accepted: 18 September 2024

Abstract

In many applied contexts, it is of interest to identify the extent to which a given association measure changes its value as different sets of variables are included in the analysis. We consider logistic regression models where the interest is for the effect of a focal binary explanatory variable on a specific response, and a further collection of binary covariates is available. We provide a methodological framework for the joint analysis of the full set of coefficients of the focal variable computed across all the models obtained by adding or removing predictors from the set of covariates. The result is obtained by applying a specific log-hybrid linear expansion of the joint distribution of the variables that implicitly comprises all the regression coefficients of interest. In this way, we obtain a method that allows one to verify, in a flexible way, a wide range of scientific hypotheses involving the comparison of multiple logistic regression coefficients both in nested and in non-nested models. The proposed methodology is illustrated through a test bed example and an empirical application.

Keywords Binary response variable · Categorical data · Contingency table · Collapsibility · Log-hybrid linear parameter · Odds ratio.

✉ Luca La Rocca
luca.larocca@unimore.it

Monia Lupparelli
monia.lupparelli@unifi.it

Alberto Roverato
alberto.roverato@unipd.it

1 Department of Physics, Informatics and Mathematics, University of Modena and Reggio Emilia, Via Campi 213/b, 41125 Modena, Italy

2 Department of Statistics, Computer Science, Applications, University of Florence, Viale Morgagni 59, 50134 Florence, Italy

3 Department of Statistical Sciences, University of Padova, Via Battisti 241, 35121 Padova, Italy

Published online: 10 October 2024

1 Introduction

In models for binary response variables, logistic regression coefficients are commonly used to measure the association between the response variable and a collection of explanatory variables; see, for instance, Chapter 5 of Agresti (2012). In this context, a lot of attention has been given to the problem of how the association between the response variable and a specific variable of interest changes its strength as different sets of explanatory variables are added or removed. Here, we attack this problem in the case where the explanatory variables are also binary.

Broadly speaking, a measure of association is said to be collapsible over a given set of variables if it does not change its value when such variables are marginalized over. Collapsibility is typically exploited to reduce dimensionality. This is especially useful when the variables are categorical and the cross-classified contingency table is sparse with a large number of sampling zeros, because collapsibility allows one to work with a lower dimensional, and thus less sparse, table. A formal approach to this problem requires to consider various definitions of collapsibility, and Guo and Geng (1995) provided conditions for both simple and strong collapsibility of logistic regression coefficients; see also Whittemore (1978), Asmussen and Edwards (1983), Wermuth (1987), and Didelez et al. (2010). Collapsibility implies the equality of parameters between nested models and, thus, a related problem concerns the comparison of parameters of different models either nested (Ducharme and Lepage 1986; Greenland and Mickey 1988; Clogg et al. 1992, 1995; Greenland et al. 1999) or non-nested (Efron 1984; Vuong 1989; Royston and Thompson 1995).

More generally, in contexts such as mediation analysis and causal inference, it is of interest to quantify how much the value of an association measure is affected or, equivalently, distorted, by the removal of some variables from the analysis (Cox and Wermuth 2003; Wermuth and Cox 2008; Xie et al. 2008). Within this framework, specific results concerning logistic regression coefficients can be found in Stanghellini and Doretti (2019), Wang (2021), Raggi et al. (2023), and references therein; see also Greenland et al. (1999) for a discussion on the distinction between non-collapsibility and confounding, with special attention to the identification of causal effects.

We consider the case where the main interest is for the effect of a focal binary variable X on a binary response Y and, furthermore, a vector $Z_M = (Z_j)_{j \in M}$ of $m = |M|$ additional binary explanatory variables has been measured. In this setting, for any given subvector $Z_D = (Z_j)_{j \in D}$ of covariates, with $D \subseteq M$, the logistic regression model with response Y and vector of predictors (X, Z_D) can be regarded as an ANOVA-like expansion of the conditional logit of Y given (X, Z_D). The coefficient β_X^D of X in this expansion encodes the effect of X on Y when the variables in Z_D take the reference, zero, level, or any other level if the interactions between X and the components of Z_D can be neglected (effect of X on Y adjusted for Z_D). Thus, depending on which covariates are selected as predictors by the choice of $D \subseteq M$, there are 2^m different logistic regression models and, accordingly, 2^m different coefficients β_X^D. For example, $\beta_X^{\emptyset}$ encodes the effect of X on Y in the regression model with no covariates (unadjusted effect) whereas β_X^M encodes the effect of X in the model including all the covariates (fully adjusted effect). Note that if $D \subseteq E$ then the model of β_X^D is nested in that of β_X^E, but if neither $D \subseteq E$ nor $E \subseteq D$ then β_X^D and β_X^E belong to non-nested models.

In this paper, we provide a framework for the joint analysis of the logistic regression coefficients β_X^D, $D \subseteq M$, which will be conveniently referred to as focal regression coefficients. Remarkably, in our approach, no logistic regression model needs to be explicitly fitted, because we rely on a specific log-hybrid linear expansion of the joint distribution of the variables (La Rocca and Roverato 2019; Roverato 2017) which results in a collection of log-hybrid linear parameters.

It is worth remarking the novelty of our approach with respect to the existing literature. The available results on collapsibility and on the comparison of parameters allow one to deal with more than one parameter for each model, but they have been developed for the comparison of two selected models. Our approach focuses on a single coefficient, but it enables us to analyze the behavior of such parameter across the full collection of models obtained by adding or removing any subset of explanatory variables. From a computational viewpoint, a single log-hybrid linear model needs to be fitted, implicitly comprising all the regression coefficients of interest. For this paper, this was effectively done using R (R Core Team 2023). Finally, although our interest lies in logistic regression coefficients, we also state our results in terms of odds ratios, because odds ratios are fundamental measures of association in categorical data analysis and, thus, they confer wider generality to our framework and facilitate the comparison with the literature.

The rest of the paper is organized as follows. The problem considered is described in Sect. 2 and then motivated by two applications in Sect. 3 (one based on a classic dataset and the second one analyzing cardiovascular disease data). In Sect. 4 we review the main features of the log-hybrid linear parameterization, whereas Sect. 5 presents the results that allow us to use the log-hybrid linear parameters to deal with the logistic regression coefficients of interest. The specific cases comprising one focal variable and up to two additional explanatory variables are detailed in Sect. 6, using the first application of Sect. 3 as an illustrative example, then Sect. 7 advances the second application of Sect. 3. Some additional technical details on the relationships existing between alternative parameterizations of binary data are provided in Appendix A, while the proofs of the results in Sect. 5 are deferred to Appendix B.

2 Preliminaries and notation

We are interested in the effect of a focal binary variable X on a binary response Y, defined as the coefficient of X in the logistic regression of Y on X and, possibly, other predictors selected from a binary vector $Z_M = (Z_j)_{j \in M}$ of covariates. If there are m covariates, we let $M = \{1, 2, \ldots, m\}$ and also write Z_M as $Z_{1:m}$, that is, we use $1\!:\!m$ as a short hand for $\{1, 2, \ldots, m\}$. Then, without loss of generality, we denote the two levels of each variable by 0 and 1, and we consider 0 as the reference level for all predictors. Finally, we assume that the joint distribution of (Y, X, Z_M) is positive, that is, $\mathrm{pr}(X = x, Y = y, Z_M = z_M) > 0$ for all $(x, y, z_M) \in \{0, 1\}^{m+2}$; note that this assumption rules out dummy variables (binary variables representing the presence or absence of each of the three or more levels of a categorical variable).

We first consider the case where no covariate is included in the analysis, so that the logistic regression model can be written as

$$\operatorname{logit} \operatorname{pr}(Y = 1 \mid X = x) \;=\; \beta_0^{\emptyset} + \beta_X^{\emptyset} x, \qquad x \in \{0, 1\}, \tag{1}$$

where $\operatorname{logit}(\pi) = \log(\pi/(1-\pi))$, $0 < \pi < 1$, while $\beta_0^{\emptyset} = \operatorname{logit} \operatorname{pr}(Y = 1 \mid X = 0)$ and $\beta_X^{\emptyset}$ is the effect of X on Y when adjusting for no covariates (as indicated by the superscript empty set symbol); see Section 5.1 of Agresti (2012). The coefficient $\beta_X^{\emptyset}$ equals the logarithm of the odds ratio, or cross-product ratio,

$$\mathrm{OR}_{XY} \;=\; \frac{\operatorname{pr}(X = 1, Y = 1)\operatorname{pr}(X = 0, Y = 0)}{\operatorname{pr}(X = 1, Y = 0)\operatorname{pr}(X = 0, Y = 1)}, \tag{2}$$

which is a well known measure of dependence for binary variables with the property that X and Y are marginally independent (in symbols $X \perp\!\!\!\perp Y$) if and only if $\mathrm{OR}_{XY} = 1$ (equivalently $\beta_X^{\emptyset} = \log \mathrm{OR}_{XY} = 0$); see Section 2.2 of Agresti (2012).

We next introduce a covariate Z_j, with $j \in M$, and assume $j = 1$, without loss of generality, so that the logistic regression model becomes (Agresti 2012, Section 5.4.1)

$$\operatorname{logit} \operatorname{pr}(Y = 1 \mid X = x, Z_1 = z_1) \;=\; \beta_0^{\{1\}} + \beta_X^{\{1\}} x + \beta_{Z_1}^{\{1\}} z_1 + \beta_{XZ_1}^{\{1\}} x z_1, \tag{3}$$

where (3) holds for all possible values of the predictors (for x and z_1 varying in $\{0, 1\}$); this latter fact is true in every logistic regression model and, thus, not explicitly stated again in the following. The logistic regression coefficients in (3) are denoted by the superscript $\{1\}$, which indicates that they are computed in the model that also includes Z_1 as a predictor: this is necessary, because (in general) the coefficients in (1) do not take the same values as those in (3), that is, both $\beta_0^{\emptyset} \neq \beta_0^{\{1\}}$ and $\beta_X^{\emptyset} \neq \beta_X^{\{1\}}$.

If we rewrite (3) by grouping the terms with and without x, we obtain

$$\operatorname{logit} \operatorname{pr}(Y = 1 \mid X = x, Z_1 = z_1) \;=\; \left(\beta_0^{\{1\}} + \beta_{Z_1}^{\{1\}} z_1\right) + \left(\beta_X^{\{1\}} + \beta_{XZ_1}^{\{1\}} z_1\right) x, \tag{4}$$

which is a version of (1) with coefficients depending on z_1: if $z_1 = 0$, they are $\beta_0^{\{1\}}$ and $\beta_X^{\{1\}}$; if $z_1 = 1$, they are $\beta_0^{\{1\}} + \beta_{Z_1}^{\{1\}}$ and $\beta_X^{\{1\}} + \beta_{XZ_1}^{\{1\}}$. Similarly, letting $j = 2$, we find

$$\operatorname{logit} \operatorname{pr}(Y = 1 \mid X = x, Z_2 = z_2) \;=\; \left(\beta_0^{\{2\}} + \beta_{Z_2}^{\{2\}} z_2\right) + \left(\beta_X^{\{2\}} + \beta_{XZ_2}^{\{2\}} z_2\right) x, \tag{5}$$

where (in general) both $\beta_0^{\{2\}} \neq \beta_0^{\{1\}}$ and $\beta_X^{\{2\}} \neq \beta_X^{\{1\}}$. Moreover, if we consider both Z_1 and Z_2 as predictors in the logistic regression model, we obtain

$$\begin{aligned} \operatorname{logit} \operatorname{pr}(Y = 1 \mid X = x, Z_{1:2} = z_{1:2}) = \left(\beta_0^{\{1,2\}} + \beta_{Z_1}^{\{1,2\}} z_1 + \beta_{Z_2}^{\{1,2\}} z_2 + \beta_{Z_1 Z_2}^{\{1,2\}} z_1 z_2\right) + \\ + \left(\beta_X^{\{1,2\}} + \beta_{XZ_1}^{\{1,2\}} z_1 + \beta_{XZ_2}^{\{1,2\}} z_2 + \beta_{XZ_1 Z_2}^{\{1,2\}} z_1 z_2\right) x, \end{aligned} \tag{6}$$

which gives rise to four distinct instances of (1) as z_1 and z_2 vary in $\{0, 1\}$.

We focus on the case where $z_1 = z_2 = 0$, so that (1), (4), (5) and (6) simplify to

$$\begin{aligned}
\operatorname{logit} \operatorname{pr}(Y = 1 \mid X = x) &= \beta_0^{\emptyset} + \beta_X^{\emptyset} x, \\
\operatorname{logit} \operatorname{pr}(Y = 1 \mid X = x, Z_1 = 0) &= \beta_0^{\{1\}} + \beta_X^{\{1\}} x, \\
\operatorname{logit} \operatorname{pr}(Y = 1 \mid X = x, Z_2 = 0) &= \beta_0^{\{2\}} + \beta_X^{\{2\}} x, \\
\operatorname{logit} \operatorname{pr}(Y = 1 \mid X = x, Z_{1:2} = 0) &= \beta_0^{\{1,2\}} + \beta_X^{\{1,2\}} x.
\end{aligned}$$

Our interest here is for the value of the coefficients $\beta_X^{\emptyset}$, $\beta_X^{\{1\}}$, $\beta_X^{\{2\}}$ and $\beta_X^{\{1,2\}}$, which represent the effect of X on Y at different depths of marginalization over Z_1 and Z_2, when conditioning on their reference levels, or any other level if Z_1 and Z_2 have null interactions with X. As formalized below in wider generality, these coefficients are conditional log odds ratios.

For every $D \subseteq M$ and $z_D \in \{0, 1\}^{|D|}$, the conditional odds ratio between X and Y given $Z_D = z_D$ can be computed as

$$\mathrm{OR}_{XY|Z_D=z_D} = \frac{\operatorname{pr}(X = 1, Y = 1 \mid Z_D = z_D) \operatorname{pr}(X = 0, Y = 0 \mid Z_D = z_D)}{\operatorname{pr}(X = 1, Y = 0 \mid Z_D = z_D) \operatorname{pr}(X = 0, Y = 1 \mid Z_D = z_D)}, \tag{7}$$

which becomes (2) when $D = \emptyset$. It is well-established that X and Y are conditionally independent given Z_D, denoted by $X \perp\!\!\!\perp Y \mid Z_D$, if and only if the conditional odds ratio (7) is equal to 1 for every $z_D \in \{0, 1\}^{|D|}$; see Section 2.3.4 of Agresti (2012).

When the covariates in Z_D are included in the analysis, for some $D \subseteq M$, we compactly write the logistic regression model, for any $E \subseteq D$, as

$$\operatorname{logit} \operatorname{pr}(Y = 1 \mid X = x, Z_E = 1, Z_{D\setminus E} = 0) = \sum_{E' \subseteq E} \beta_{Z_{E'}}^D + \sum_{E' \subseteq E} \beta_{XZ_{E'}}^D x, \tag{8}$$

where $Z_E = 1$ means that $Z_i = 1$ for every $i \in E$ (similarly for $Z_{D\setminus E} = 0$) and we use the convention $\beta_{Z_\emptyset}^D = \beta_0^D$. For $E = \emptyset$, the equality in (8) simplifies to

$$\operatorname{logit} \operatorname{pr}(Y = 1 \mid X = x, Z_D = 0) = \beta_0^D + \beta_X^D x,$$

whereas, for $E \neq \emptyset$, if $\beta_{XZ_{E'}}^D = 0$ for all non-empty $E' \subseteq E$, it takes the form

$$\operatorname{logit} \operatorname{pr}(Y = 1 \mid X = x, Z_E = 1, Z_{D\setminus E} = 0) = \sum_{E' \subseteq E} \beta_{Z_{E'}}^D + \beta_X^D x,$$

which covers the case where X and the components of Z_M have null interactions. Our interest is for the relationship existing among the coefficients β_X^D, $D \subseteq M$.

The coefficients β_X^D, $D \subseteq M$, are conditional log odds ratio, according to the following identity:

$$\beta_X^D = \log \mathrm{OR}_{XY|Z_D=0} \quad \text{for every } D \subseteq M. \tag{9}$$

The identity (9) is well-known and given, among others, in Sections 5.3 and 9.5 of Agresti (2012). However, it represents a central point of our work and therefore, for the sake of completeness, we formally derive it as (A8) in Appendix A, where we also provide a more formal review of logistic regression, including a formula to compute logistic regression coefficients from cell probabilities.

3 Motivating examples

This section introduces two motivating examples to emphasize the interest in comparing logistic regression coefficients across different models.

3.1 Titanic data

The first example we consider is a set of data on the tragic sinking of Titanic in 1912, where there is an interest in studying the impact that the travelling class had on the survival of passengers. It can be regarded as a classic, which makes it ideal for illustrative purposes, and it exists in several versions, of which we consider the one known as *titanic3* (obtained from http://hbiostat.org/data courtesy of the Vanderbilt University Department of Biostatistics). For instance, the same set of data was recently analyzed by Yin et al. (2022) using models based on relative risk parameters. It consists of four variables collected on 1046 passengers: *sur*, survival (0: no, 1: yes); *class*, travelling class (0: first or second, 1: third); *sex* (0: male, 1: female); *age* (0: $\geq$ 25 years, 1: $<$ 25 years). Note that we dichotomized age and $class$, and coded the variables so that the coefficient of *class* (the focal variable) will measure the association between travelling in third class and survival ($sur = 1$) for elderly males (when including age and sex as predictors).

We consider four different logistic regression models, obtained by combining the reference levels of *sex* and *age*:

$$\operatorname{logit} \operatorname{pr}(sur = \text{yes} \mid class) = \beta_0^{\emptyset} + \beta^{\emptyset}\, \text{class} \tag{10}$$
$$\operatorname{logit} \operatorname{pr}(sur = \text{yes} \mid class, \text{male}) = \beta_0^{\text{sex}} + \beta^{\text{sex}}\, \text{class} \tag{11}$$
$$\operatorname{logit} \operatorname{pr}(sur = \text{yes} \mid class, \geq 25) = \beta_0^{\text{age}} + \beta^{\text{age}}\, \text{class} \tag{12}$$
$$\operatorname{logit} \operatorname{pr}(sur = \text{yes} \mid class, \text{male}, \geq 25) = \beta_0^{\text{sex,age}} + \beta^{\text{sex,age}}\, \text{class} \tag{13}$$

where the focal coefficient β^D in Eq. (10)–(13) represents the effect of $class$ on sur at the depth of marginalisation over sex and age specified by $D \subseteq \{sex, age\}$; note that, for the sake of readability, we dropped the subscript $class$ from β_{class}^D and the curly brackets in the specification of D.

Table 1 Titanic data. MLE (ASE) of the four focal coefficients

$\beta^{\emptyset}$	β^{sex}	β^{age}	$\beta^{sex,age}$
−1.211 (0.133)	−0.472 (0.194)	−1.270 (0.186)	−0.331 (0.254)

The Maximum Likelihood Estimate (MLE) of the focal coefficient β^D is given in Table 1, together with its Asymptotic Standard Error (ASE) from the observed Fisher information matrix, for the four different choices of D (models); note that the MLE of β^D equals the log odds ratio between *class* and *sur* computed from their conditional proportions given the variables in D take their reference level. The estimated effect of travelling in third class on survival is always negative, but its value varies considerably across models, in particular when *sex* is included as a predictor. In Sect. 6.3, we will use log-hybrid linear parameters to carry out a comparison between the four focal coefficients of this example.

3.2 National march cohort data

Our second example represents a case where a partition of the set of covariates is appropriate. Indeed, sometimes, we can distinguish between pure context variables, or background variables, and intermediate variables that are explanatory with respect to the final outcome, but might be responses with respect to the focal variable. Mediation analysis is a relevant example, where the focal variable represents a treatment, or an exposure, the background variables are known as confounders, and the intermediate variables, called mediators, are able to capture and channel the indirect effect of the treatment, or exposure, on the final outcome (VanderWeele 2015).

As a representative example, we consider a set of data from the observational Swedish National March Cohort (NMC): a population of 33, 327 individuals interviewed in 1997, on the occasion of a fund raising event, and followed from that year to 2004, through the Swedish patient registry (Sjölander et al. 2009). The NMC data comprise six variables: *pac*, self-reported physical activity level (0: high-level exerciser, 1: low-level exerciser) considered as a treatment; *bmi*, body mass index (0: ≥ 30, 1: < 30) and *sle*, sleeping problems (0: yes, 1: no) considered as mediators or intermediate variables; *sex* (0: male, 1: female) and *age* (0: ≥ 50, 1: < 50) considered as pure context variables; *cvd*, cardiovascular disease events (0: no event, 1: at least one event) considered as the final outcome, or response variable. Variables *bmi*, *sle*, *age* and *sex* have been dichotomized so that level 0 identifies subjects with adverse conditions for cardiovascular diseases. Conversely, the adverse condition of variable *pac* corresponds to level 1, so that its regression coefficient will measure the association between a lack of physical activity and the insurgence of cardiovascular disease events.

Low-level exercisers are expected to have a higher risk of cardiovascular diseases both for a direct adverse effect given by the lack of physical activity and for an indirect adverse effect given by the worsening that the lack of physical activity can give in terms of body mass index and sleeping problems. Then, the main interest in this application becomes twofold: (i) studying the overall effect of physical activity (pac) on the adverse outcome (cvd) regardless of information on bmi and sle; (ii) distinguishing between the direct effect of pac and the indirect effect channelled by bmi and sle (Pearl 2001; Schwartz et al. 2011). This should be done considering variables age and sex as confounders, useful to adjust the treatment effect to the population of interest (and make it amenable to causal interpretation).

We therefore consider regression models at different depths of marginalization over the intermediate variables bmi and sle, but always including sex and age as predictors. In these models, all interaction terms, in particular those involving pac, are non-significant, so that the effect of physical activity can be interpreted regardless of the value of the other explanatory variables. Nevertheless, to fix ideas, we write:

$$\operatorname{logit}\operatorname{pr}(cvd = 1 \mid pac, \text{male}, \geq 50) = \beta_0^* + \beta^*\, pac \tag{14}$$

$$\operatorname{logit}\operatorname{pr}(cvd = 1 \mid pac, \text{male}, \geq 50, bmi \geq 30) = \beta_0^{*,bmi} + \beta^{*,bmi}\, pac \tag{15}$$

$$\operatorname{logit}\operatorname{pr}(cvd = 1 \mid pac, \text{male}, \geq 50, sle = \text{yes}) = \beta_0^{*,sle} + \beta^{*,sle}\, pac \tag{16}$$

$$\operatorname{logit}\operatorname{pr}(cvd = 1 \mid pac, \text{male}, \geq 50, bmi \geq 30, sle = \text{yes}) = \beta_0^{*,bmi,sle} + \beta^{*,bmi,sle} pac \tag{17}$$

where * stands for sex, age (always included as predictors with reference level for elderly male). The focal coefficient $\beta^* = \beta^{sex,age}$ in (14) represents the overall effect of pac on cvd, adjusted for sex and age, while the same coefficient in the other regression equations can be interpreted as the direct effect of pac on cvd for elderly males with high body mass index (15), sleeping problems (16) and both of them (17).

The differences between the direct effects in Eqs. (15)–(17) and the overall effect in equation (14) account both for the indirect effects that flow through the two mediators (individually and as a whole) and for the non-collapsibility of odds ratios, as discussed by Raggi et al. (2023); one should not expect these differences to vanish if the mediators are independent of the focal variable. With this limitation in mind, regardless of the causal model specification and related interpretation, Sect. 7 shows how the log-hybrid parameterization can be employed to decompose the treatment effect with respect to the inclusion of mediators.

4 The log-hybrid linear parameterization

The framework proposed in this paper to enable the comparison of logistic regression coefficients across different models, that is, conditional odds ratios at different depths of marginalization, relies on expressing the quantities of interest in terms of a parameterization, called the log-hybrid linear parameterization, introduced by La Rocca and Roverato (2019). We recall in this section the essential features of this parameterization, and we refer to Appendix A for additional details, including a formula to compute log-hybrid linear parameters from cell probabilities.

In the log-hybrid linear parameterization, the set of variables $\{X, Y\} \cup \{Z_i\}_{i \in M}$ needs to be partitioned into two disjoint subsets: a set of core variables that should always be included in the analysis, and a set of peripheral variables that may or may not be included in the analysis. This partition affects the interpretation of the parameters and, to start with, we partition the variables into $\{X, Y\}$ and $\{Z_i\}_{i\in M}$. With this choice, using the notation of Appendix A, the log-hybrid linear parameters can be written as $\varphi_{A\cup D}^{\{X,Y\}}$ with $A \subseteq \{X, Y\}$ and $D \subseteq \{Z_i\}_{i\in M}$. In fact, since we do not need the full set of parameters, but only those with $A = \{X, Y\}$, we can simplify the notation and denote the collection of parameters of interest as $\varphi_{XYZ_D}^{XY}$, $D \subseteq M$.

In greater generality, we partition the variables into $\{X, Y\}\cup\{Z_i\}_{i\in C}$ and $\{Z_i\}_{i\in M\setminus C}$, where $C \subset M$. In this way, the log-hybrid linear parameters we consider can be written as $\varphi_{A\cup B\cup D}^{\{X,Y\}\cup\{Z_i\}_{i\in C}}$ with $A \subseteq \{X, Y\}$, $B \subseteq \{Z_i\}_{i\in C}$ and $D \subseteq \{Z_i\}_{i\in M\setminus C}$. Furthermore, since we are only interested in the parameters with $A = \{X, Y\}$ and $B = \emptyset$, we can denote the collection of parameters of interest as

$$\varphi_{XYZ_D}^{XYZ_C}, \qquad D \subseteq M \setminus C.$$

Clearly, letting $C = \emptyset$, we recover the parameters $\varphi_{XYZ_D}^{XY}$, $D \subseteq M$ as a special case.

La Rocca and Roverato (2019) have shown that there exists a one-to-one relationship between the log-hybrid linear parameters and the cell probabilities of (X, Y, Z_M). More specifically, such relationship is a smooth bijection and, thus, the log-hybrid linear parameterization is a valid reparameterization of the joint distribution of (X, Y, Z_M). This implies that reliable algorithms for the computation of MLEs are available (Lang 1996) under linear constraints on the log-hybrid linear parameters. The corresponding *log-hybrid linear models* are useful both to achieve parametric parsimony, by setting higher-order interactions to zero, and to express the equality of focal coefficients, using the results in Sect. 5. Furthermore, standard asymptotic theory can be applied so that, for instance, the asymptotic distribution of MLEs is normal and likelihood ratios are asymptotically chi-squared distributed; see also Chapter 6 of Roverato (2017).

The reason why the log-hybrid linear parameterization is germane to the setting we consider is that, for every $D \subseteq M\setminus C$, the parameter $\varphi_{XYZ_D}^{XYZ_C}$ is computed from the distribution of (X, Y, Z_C, Z_D). For instance, letting for simplicity $C = \emptyset$, the parameter with $D = \emptyset$, denoted by φ_{XY}^{XY}, is computed from the distribution of (X, Y), while the parameter with $D = M$, denoted by $\varphi_{XYZ_M}^{XY}$, is computed from the distribution of (X, Y, Z_M). Hence, different log-hybrid linear parameters are associated to distributions at different depths of marginalization over Z_M and thus, as shown in the next section, to logistic regression coefficients in models including different, possibly non-nested, sets of covariates.

We remark that, for categorical variables, parameter values typically depend on the way variables are coded. This is also true for the log-hybrid linear parameters, which may change their value if the 0 and 1 labels are swapped. Hence, if we denote by $\bar{Z}_D = 1 - Z_D$ the label swapped version of Z_D, for a given $D \subseteq M$, so that $\mathrm{pr}(Z_j = 1) = \mathrm{pr}(\bar{Z}_j = 0)$, for every $j \in D$, then the parameters $\varphi_{XYZ_D}^{XY}$ and $\varphi_{XY\bar{Z}_D}^{XY}$ will typically not be equal. On the other hand, the log-hybrid linear parameters were originally introduced to parameterize regression graph models (Cox and Wermuth 1996; Drton 2009) because there is a one-to-one correspondence between the independencies implied by this family of graphical models and the vanishing of certain log-hybrid linear parameters: this correspondence holds no matter the way variables are coded.

5 Links between different depths of marginalization

We show in this section that each focal parameter β_X^D, with $D \subseteq M$, can be written as a linear combination of certain log-hybrid linear parameters. In this way, we obtain an explicit representation of the relationship existing between β_X^D and the "nested" coefficients $\beta_X^{D'}$, $D' \subset D$. Furthermore, any linear combination of regression coefficients can be written as a linear combination of log-hybrid linear parameters for which standard inferential procedures are available. A relevant example is provided by the difference $\beta_X^E - \beta_X^D$, with $D, E \subseteq M$, which enables the comparison of any two focal regression coefficients belonging to either nested or non-nested models.

As a refinement of our analysis, we restrict our attention to the coefficients $\beta_X^{C \cup D}$, $D \subseteq M \backslash C$, for a certain $C \subset M$. In this way, we tailor our framework to the case where the components of Z_C should always be included in the model, whereas the components of $Z_{M \backslash C}$ may or may not be. This case is crucial to mediation analysis: if C is a minimal sufficient adjustment set, while $M \setminus C$ is a set of candidate mediators, the vector Z_C should always be conditioned upon, in order for the effect of X on Y to admit a causal interpretation, while the vector Z_D, with $D \subseteq M \backslash C$, should only be conditioned upon to compute the corresponding direct effect, provided that identifying conditions are met; see for instance Pearl (2012). We call the model specified by $D = \emptyset$ the base model (representing the minimal model we are interested in) and note that letting $C = \emptyset$ returns the unrefined analysis.

We state below three results linking logistic regression coefficients at different depths of marginalization and log-hybrid linear parameters with different sets of core variables. The proofs of these results can be found in Appendix B, while their practical application to data analysis will be illustrated in the next two sections. The starting point is the following basic result concerning the log-hybrid linear parameterization.

Lemma 1 *Let X, Y and the components of $Z_M = (Z_i)_{i \in M}$ be binary random variables with positive joint probability distribution. Then, for every $D \subseteq M$, it holds that*

$$\varphi_{XY}^{XY|Z_D=1} = \sum_{D' \subseteq D} \varphi_{XYZ_{D'}}^{XY},$$

where $\varphi_{XY}^{XY|Z_D=1}$ is the log-hybrid linear parameter φ_{XY}^{XY} computed from the conditional probability distribution of X and Y given $Z_D = 1$ (their marginal distribution when $D = \emptyset$).

Proof See Appendix B. □

Lemma 1 leads to the following theorem, which links logistic regression coefficients and log-hybrid linear parameters. Recall that β_X^D is the coefficient of X in the logistic regression model with Y as response variable and the components of Z_D, together with X, as predictors.

Theorem 1 *In the setting of Lemma 1, for every $D \subseteq M$, it holds that*

$$\beta_X^D = \log \mathrm{OR}_{XY|Z_D=0} = \sum_{D' \subseteq D} \varphi^{XY}_{XY\bar{Z}_{D'}}, \tag{18}$$

where $\bar{Z}_{D'}$ is the label swapped version of $Z_{D'}$ (for all $D' \subseteq D$).

Proof See Appendix B. □

The relevance of Theorem 1 stems from the fact that it provides a way to carry out joint statistical inference on the coefficients of X in different, possibly non-nested, logistic regression models. We clarify this issue by means of an example. Let D and E be two distinct, but otherwise arbitrary, subsets of M. Then, the corresponding coefficients β_X^D and β_X^E belong to the logistic regression models with the components of Z_D and Z_E, respectively, as additional predictors. The two models are nested if either $D \subset E$ or $E \subset D$, but they are non-nested otherwise. Now, Eq. (18) establishes a linear relationship between the coefficients β_X^D, β_X^E and the terms $\varphi^{XY}_{XY\bar{Z}_F}$, $F \subseteq D \cup E \subseteq M$, which implies that the difference $\beta_X^E - \beta_X^D$ can be written as a linear combination of φ-terms. Since the φ-terms in (18) belong to the log-hybrid linear expansion of the cell probabilities of $(X, Y, \bar{Z}_M)$, we can apply standard inferential methods to carry out both estimation and testing on $\beta_X^E - \beta_X^D$. The same holds true for any other linear combination of focal coefficients, because it can be immediately expressed as a linear combination of φ-terms.

More specifically, from standard inferential methods we can compute the MLEs of the parameters of the relevant log-hybrid linear model and then obtain their joint asymptotic normal distribution. From the latter, we can exploit Theorem 1 (or Theorem 3 below) to compute the estimate of any linear combination of focal coefficients, together with its asymptotic normal distribution.

Another implication of (18) is that any coefficient β_X^D can be written as a linear combination of the corresponding coefficients in all nested logistic regression models, with the addition of the term $\varphi^{XY}_{XY\bar{Z}_D}$. We state this fact in the following result, whose application and interpretation will be clarified in Sect. 6, where the cases with one and two covariates will be analyzed in detail.

Theorem 2 *In the setting of Lemma 1, for every $D \subseteq M$, it holds that*

$$\beta_X^D = \sum_{D' \subset D} (-1)^{|D \setminus D'|-1} \beta_X^{D'} + \varphi^{XY}_{XY\bar{Z}_D}, \tag{19}$$

or, equivalently,

$$\log \mathrm{OR}_{XY|Z_D=0} = \sum_{D' \subset D} (-1)^{|D \setminus D'|-1} \log \mathrm{OR}_{XY|Z_{D'}=0} + \varphi^{XY}_{XY\bar{Z}_D}. \tag{20}$$

where $\bar{Z}_D$ is the label swapped version of Z_D.

Proof See Appendix B. □

The φ-terms in (18) are computed with X and Y as the sole core variables. If the components of a vector of covariates Z_C, with $C \subseteq M$, are added to the core variables, a generalization of (18) still holds, which leads to a direct and parsimonious link between β_C and $\beta_{C \cup D}$ for $D \subseteq M \setminus C$. We state this link in the following result, which will be applied and discussed in Sect. 7, and remark that for $C = \emptyset$ it coincides with (18).

Theorem 3 *In the setting of Lemma 1, let $C \subset M$. Then, for every $D \subseteq M \setminus C$, it holds that $\beta_X^C = \log \mathrm{OR}_{XY|Z_C=0} = \varphi_{XY}^{XYZ_C}$ and*

$$\beta_X^{C \cup D} = \beta_X^C + \sum_{\emptyset \neq D' \subseteq D} \varphi_{XY\bar{Z}_{D'}}^{XYZ_C}, \tag{21}$$

or, equivalently,

$$\log \mathrm{OR}_{XY|Z_C=0, Z_D=0} = \log \mathrm{OR}_{XY|Z_C=0} + \sum_{\emptyset \neq D' \subseteq D} \varphi_{XY\bar{Z}_{D'}}^{XYZ_C}, \tag{22}$$

where $\bar{Z}_{D'}$ is the label swapped version of $Z_{D'}$ (for all $D' \subseteq D$).

Proof See Appendix B. □

Although our main interest is for logistic regression coefficients, in Theorem 1, as well as in Theorem 2 and Theorem 3, the results are also given in terms of conditional log odds ratios. This choice is motivated by the fact that odds ratios are fundamental measures of association in categorical data analysis, and thus they facilitate both the comparison with the literature and the extension of our results to other parameterizations, such as the log-linear parameterization; see Appendix A. In this perspective, we like to recall that (functions of) the odds ratios are the only measures of association that do not depend on the marginal distributions (Edwards 1963; Altham 1970) and the relationship between conditional independence and vanishing conditional log odds ratios is straightforward.

Finally, we find it useful to relate the material described in this section with the collapsibility of odds ratios, and thus of logistic regression coefficients. It is well-established that, for every non-empty $D \subseteq M$, a sufficient condition for the equality $\beta_X^D = \beta_X^{\emptyset}$ to hold true is that either (i) $Y \perp\!\!\!\perp Z_D \mid X$ or (ii) $X \perp\!\!\!\perp Z_D \mid Y$ (Whittemore 1978; Guo and Geng 1995; Didelez et al. 2010). Additional insight can be obtained from Theorem 3 of La Rocca and Roverato (2019), where it is shown that conditions (i) and (ii) above are both sufficient to have $\varphi_{XY\bar{Z}_{D'}}^{XY} = 0$ for every non-empty $D' \subseteq D$. In other words, with the exception of φ_{XY}^{XY}, all the remaining φ-terms in (18) are equal to zero; see also Theorem 6.3 in Roverato (2017). In this way, when either (i) or (ii) hold, Eq. (18) simplifies to $\beta_X^D = \varphi_{XY}^{XY} = \beta_X^{\emptyset}$, which returns the well-known result.

As shown in Sect. 6.2 below, the results on collapsibility can be exploited to simplify the relationships among parameters given in Eq. (19) and (20) of Theorem 2. Moreover, additional simplifications not implied by collapsibility are possible, because La Rocca and Roverato (2019) and Roverato (2017), in Theorem 6.5, have shown that a sufficient condition for $\varphi^{XY}_{XY\bar{Z}_D}$ to vanish is that there exists a partition $D_1 \cup D_2 = D$, with $D_1 \cap D_2 = \emptyset$, such that Z_{D_1} and Z_{D_2} are conditionally independent given X and Y. More specifically, if $Z_{D_1} \perp\!\!\!\perp Z_{D_2} \mid (X, Y)$, then $\varphi^{XY}_{XY\bar{Z}_{D'}} = 0$ for every $D' \subseteq D$ such that both $D' \cap D_1 \neq \emptyset$ and $D' \cap D_2 \neq \emptyset$.

6 Specific cases with up to two covariates

In this section we illustrate the application of the results in Sect. 5 to the specific cases of models including the focal variable X and either one or two additional explanatory variables.

6.1 One covariate case

We illustrate here the case where only one additional explanatory variable Z_1 is considered and, thus, we compare the logistic regression model in (1) with that given in (4). We remark that this is, in fact, a well-understood problem that has been investigated from different viewpoints. For instance, conditions for the equality $\beta^{\{1\}}_X = \beta^{\emptyset}_X$ to hold true follow from the results on collapsibility of Guo and Geng (1995) and Whittemore (1978), whereas a statistical test on hypotheses concerning the difference $\beta^{\{1\}}_X - \beta^{\emptyset}_X$ can be found in Ducharme and Lepage (1986).

Here, we notice that when $D = \{1\}$ Eq. (19) of Theorem 2 becomes

$$\beta^{\{1\}}_X = \beta^{\emptyset}_X + \varphi^{XY}_{XY\bar{Z}_1}, \tag{23}$$

thereby highlighting the role played by the log-hybrid linear parameters in this case. Clearly, $\beta^{\{1\}}_X - \beta^{\emptyset}_X = \varphi^{XY}_{XY\bar{Z}_1}$, so that $\beta^{\{1\}}_X = \beta^{\emptyset}_X$ if and only if $\varphi^{XY}_{XY\bar{Z}_1} = 0$. Hence, the comparison of the two focal coefficients can be based on the MLE $\widehat{\varphi}^{XY}_{XY\bar{Z}_1}$ of $\varphi^{XY}_{XY\bar{Z}_1}$. Furthermore, as clarified by Sect. 6.2 below, this interpretation of the parameter $\varphi^{XY}_{XY\bar{Z}_1}$ remains unchanged when additional explanatory variables are introduced in the model, due to the upper compatibility property of the log-hybrid parameterization.

Finally, we recall that Eq. (23) is closely related to Eq. (5) of Stanghellini and Doretti (2019), where $\beta^{\emptyset}_X$ is written as the sum of $\beta^{\{1\}}_X$ and other three terms, which can thus be regarded as a decomposition of $-\varphi^{XY}_{XY\bar{Z}_1}$; see also Wang (2021). The relevance of such a decomposition has to do with the interpretation of the relationship existing between the two models, because it involves only terms that are computed on the joint distribution of the three variables, without requiring the computation of the marginal distribution of X and Y.

6.2 Two covariate case

In the case where $D = \{1, 2\}$, Eq. (19) of Theorem 2 becomes

$$\beta_X^{\{1,2\}} = -\beta_X^{\emptyset} + \beta_X^{\{1\}} + \beta_X^{\{2\}} + \varphi_{XY\bar{Z}_1\bar{Z}_2}^{XY}, \tag{24}$$

which provides an expansion of the coefficient $\beta_X^{\{1,2\}}$ of (6) into (i) the coefficient $\beta_X^{\emptyset}$ from (1), computed on the marginal distribution of X and Y, (ii) the coefficient $\beta_X^{\{1\}}$ from (4), computed on the marginal distribution of X, Y and Z_1, (iii) the coefficient $\beta_X^{\{2\}}$ from (5), computed on the marginal distribution of X, Y and Z_2, and finally (iv) the term $\varphi_{XY\bar{Z}_1\bar{Z}_2}^{XY}$ computed on the full joint distribution of the four variables.

The expansion of $\beta_X^{\{1,2\}}$ in (24), properly combined with (18) from Theorem 1, can be exploited to verify a wide range of scientific hypotheses, in a flexible way. For example, one can compare the focal coefficients of both nested models, by means of a difference such as

$$\beta_X^{\{1,2\}} - \beta_X^{\{2\}} = \varphi_{XY\bar{Z}_1}^{XY} + \varphi_{XY\bar{Z}_1\bar{Z}_2}^{XY},$$

and non-nested models, by means of a difference such as

$$\beta_X^{\{1\}} - \beta_X^{\{2\}} = \varphi_{XY\bar{Z}_1}^{XY} - \varphi_{XY\bar{Z}_2}^{XY}.$$

Furthermore, if $\varphi_{XY\bar{Z}_1\bar{Z}_2}^{XY} = 0$ then (24) simplifies to

$$\beta_X^{\{1,2\}} = -\beta_X^{\emptyset} + \beta_X^{\{1\}} + \beta_X^{\{2\}}, \tag{25}$$

which shows how the coefficients of smaller logistic regression models contribute to form the value of $\beta_X^{\{1,2\}}$. It is worth remarking that this equation provides a relationship between logistic regression coefficients that cannot be obtained from the application of the results on collapsibility, and that implies, for example, both

$$\beta_X^{\emptyset} - \beta_X^{\{1\}} = \beta_X^{\{2\}} - \beta_X^{\{1,2\}} \quad \text{and} \quad \beta_X^{\emptyset} - \beta_X^{\{2\}} = \beta_X^{\{1\}} - \beta_X^{\{1,2\}}; \tag{26}$$

that is, removing Z_1 from the model has the same impact on the coefficient of X whether Z_2 is present or not in the model and, symmetrically, the same is true if Z_2 is removed instead. In addition, if $\varphi_{XY\bar{Z}_1}^{XY}$ also vanishes, we see from (23) that $\beta_X^{\emptyset} = \beta_X^{\{1\}}$, and thus (25) becomes

$$\beta_X^{\{1,2\}} = \beta_X^{\{2\}}.$$

Finally, we recall that the equality $\varphi_{XY\bar{Z}_1\bar{Z}_2}^{XY} = 0$ is implied by any of the conditions $Z_1 \perp\!\!\!\perp Z_2 \mid (X, Y)$, $X \perp\!\!\!\perp Z_{1:2} \mid Y$ and $Y \perp\!\!\!\perp Z_{1:2} \mid X$, whereas the equality $\varphi_{XY\bar{Z}_1}^{XY} = 0$

Table 2 Titanic data. Saturated model. MLE (ASE) of the log-hybrid linear parameters of interest and related focal coefficients

$\varphi_{sur,class}$	$\varphi_{sur,class,\overline{sex}}$	$\varphi_{sur,class,\overline{age}}$	$\varphi_{sur,class,\overline{sex},\overline{age}}$
−1.211 (0.133)	0.739 (0.142)	−0.059 (0.130)	0.200 (0.120)
$\beta^{\emptyset}$	β^{sex}	β^{age}	$\beta^{sex,age}$
−1.211 (0.133)	−0.472 (0.194)	−1.270 (0.186)	−0.331 (0.254)

is implied by any of the conditions $X \perp\!\!\!\perp Z_1 \mid Y$ and $Y \perp\!\!\!\perp Z_1 \mid X$. However, these are sufficient conditions, and the vanishing of any φ-term, as well as of any linear combination of φ-terms, can be directly verified with standard inferential procedures relying on the asymptotic normal distribution of the MLEs of the log-hybrid linear parameters.

6.3 Illustrative example with the Titanic data

The log-hybrid linear parameters with core variables $class$ and sur are estimated without constraints from the joint distribution of the variables sur, $class$, sex and age of the Titanic data introduced in Section 3.1; the results are reported in Table 2. Note that sex and age are label swapped, before estimation, so that the results in Sect. 5 can be applied. Indeed, Table 2 also includes the estimated coefficients in Eqs. (10)–(13) obtained from Theorem 1:

$$\begin{aligned}
\hat{\beta}^{\emptyset} &= \hat{\varphi}_{sur,class}, \\
\hat{\beta}^{sex} &= \hat{\varphi}_{sur,class} + \hat{\varphi}_{sur,class,\overline{sex}}, \\
\hat{\beta}^{age} &= \hat{\varphi}_{sur,class} + \hat{\varphi}_{sur,class,\overline{age}}, \\
\hat{\beta}^{sex,age} &= \hat{\varphi}_{sur,class} + \hat{\varphi}_{sur,class,\overline{sex}} + \hat{\varphi}_{sur,class,\overline{age}} + \hat{\varphi}_{sur,class,\overline{sex},\overline{age}},
\end{aligned}$$

where the hat denotes MLEs and, for the sake of readability, we dropped the superscript $sur, class$ common to all φ-terms. Recall that the above displayed β-terms are the focal regression coefficients at different depths of marginalization over age and sex, which represent the quantities of interest in our analysis.

Single log-hybrid linear parameters or sums of them can be interpreted as covariate-specific modifiers of the focal coefficient of the base model. Parameter $\varphi_{sur,class,\overline{sex}}$ is defined as the first order sex-specific modifier: it is significant, and thus the contribution of variable sex is shown to be important to this extent. Similarly, $\varphi_{sur,class,\overline{age}}$ is the first order age-specific modifier: it is non-significant, and thus variable age seems to be negligible as a modifier of the focal parameter. Let us now consider the second order modifier defined as

$$\hat{\beta}^{sex,age} - \hat{\beta}^{\emptyset} = \hat{\varphi}_{sur,class,\overline{sex}} + \hat{\varphi}_{sur,class,\overline{age}} + \hat{\varphi}_{sur,class,\overline{sex},\overline{age}},$$

Table 3 Titanic data. Reduced model. MLE (ASE) of the log-hybrid linear parameters of interest and related focal coefficients

$\varphi_{sur,class}$	$\varphi_{sur,class,\overline{sex}}$	$\varphi_{sur,class,\overline{age}}$	$\varphi_{sur,class,\overline{sex},\overline{age}}$
−1.211 (0.133)	0.755 (0.141)	−0.042 (0.129)	0
$\beta^{\emptyset}$	β^{sex}	β^{age}	$\beta^{sex,age}$
−1.211 (0.133)	−0.456 (0.193)	−1.253 (0.185)	−0.498 (0.238)

related to the inclusion of both variables sex and age. The estimate of this modifier accounting for both sex and age is 0.880 (0.216) and thus, it is significant.

Using a stepwise procedure, based on the hierarchical principle, we verify whether the saturated log-hybrid model can be simplified by the inclusion of zero constraints. A reduced model with $\varphi_{sur,class,\overline{sex},\overline{age}} = 0$ is fitted, since this is the highest non-significant interaction term. The resulting model provides an acceptable fit (likelihood ratio test statistic equal to 2.69 on 1 degree of freedom, p-value = 0.10) and the parameter estimates are included in Table 3. The immediate simplification is that the second order modifier is additive:

$$\hat{\beta}^{sex,age} - \hat{\beta}^{\emptyset} = \hat{\varphi}_{sur,class,\overline{sex}} + \hat{\varphi}_{sur,class,\overline{age}}, \tag{27}$$

that is, it is the sum of the first order modifiers for sex and age; its estimate is 0.713 (0.197) and thus it remains significant.

As shown in (25), the simplification given by (27) provides a nice interpretation of the focal parameter estimate of the model including both sex and age: it becomes a linear combination of the focal parameter estimates in the two nested models,

$$\hat{\beta}^{sex,age} = -\hat{\beta}^{\emptyset} + \hat{\beta}^{sex} + \hat{\beta}^{age},$$

so that both $\hat{\beta}^{\emptyset} - \hat{\beta}^{sex} = \hat{\beta}^{age} - \hat{\beta}^{sex,age}$ and $\hat{\beta}^{\emptyset} - \hat{\beta}^{age} = \hat{\beta}^{sex} - \hat{\beta}^{sex,age}$, similarly to (26), which means that removing sex or age from the linear predictor has the same impact on the focal parameter regardless of the depth of marginalization.

In the reduced model, the greatest contribution to the focal parameter still comes from the inclusion of sex, whereas $\varphi_{sur,class,\overline{age}}$ again results to be non-significant. Then, we procede with the stepwise procedure and compare the reduced model defined by $\varphi_{sur,class,\overline{age}} = \varphi_{sur,class,\overline{sex},\overline{age}} = 0$ with the saturated model. We remark that such a comparison is equivalent to testing the joint hypothesis

$$\beta^{age} = \beta^{\emptyset} \quad \text{and} \quad \beta^{sex,age} = \beta^{sex}, \tag{28}$$

which expresses the null impact of removing age on the focal parameter at any depth of marginalization. In this specific case, the empirical evidence is against such a simplification (likelihood ratio test statistic equal to 28.21 on 2 degree of freedom, p-value ≈ 0). Note that (28) simultaneously involves the coefficients of four different logistic regression models and, when verified, would represent a weaker assumption than collapsibility of the focal coefficient with respect to age.

We conclude that, despite the age-specific modifier $\varphi_{sur,class,\overline{age}}$ seems to be negligible in terms of its impact on the estimate of the focal parameter, the model with $\varphi_{sur,class,\overline{age}} = \varphi_{sur,class,\overline{sex},\overline{age}} = 0$ does not provide an adequate fit, and thus the model with $\varphi_{sur,class,\overline{sex},\overline{age}} = 0$ is selected. Form Table 3, one can see that the greatest impact on the focal parameter is provided by the positive estimate of the sex-specific modifier $\varphi_{sur,class,\overline{sex}}$, which considerably reduces the negative impact of travelling in third class on the survival probability for male passengers. Conversely, including covariate age in the predictor slightly increases the negative effect of the travelling class on the survival probability for passengers over 25 years old. When both covariates sex and age are included, the estimate of their second order modifier in Eq. (27) is additive and obtained by summing the first order modifiers that are, respectively, positive and negative for sex and age. The resulting sum is positive and can be interpreted as an attenuation of the inclusion of sex on the effect of travelling in third class for male passengers if they are over 25 years old.

In the end, positive (negative) log-hybrid linear parameter estimates, as well as positive (negative) sums of some of them, increase (reduce) the effect of the focal variable on the probability of success of the response at different depths of marginalization over the covariates, with reference to their baseline level; in this specific application with reference to sex = male and $age \geq 25$.

This illustrative example was discussed with a dual purpose. Firstly, to show that the log-hybrid linear parameterization effectively enables the comparison of coefficients across different logistic regression models. Secondly, to illustrate how zero constraints on log-hybrid linear parameters can be used to test hypotheses of interest on the focal parameter at different depths of marginalization.

7 Application to NMC data

In this section a log-hybrid linear model is applied for the analysis of the NMC data introduced in Sect. 3.2. The interest lies in studying the effect of a treatment, physical activity (pac), on the insurgence of cardiovascular disease events (cvd), by considering biological sex (sex) and age (age) of the respondents as confounders, and their body mass index (bmi) and sleeping problems (sle) as possible mediators. In a logistic regression setting, the focal parameter is represented by the coefficient of pac and the intent is to study how this parameter can be influenced by the inclusion of bmi and/or sle in the linear predictor. The substantive hypothesis is that physical activity has a twofold effect on the risk of cardiovascular disease events: a direct effect and an indirect effect induced by the increase of body mass index and sleeping problems in low level exercisers.

The above question can be approached by fitting a log-hybrid linear model to the joint distribution of the full set $\{cvd, pac, age, sex, \overline{bmi}, \overline{sle}\}$ of variables and by exploiting the results provided in this paper to decompose and compare coefficients in multiple logistic regression models. In this analysis, the base model includes the treatment pac and the confounders sex and age, while the variables bmi and sle play the role of mediators, and thus they are used with swapped labels (as required by our results in Sect. 5). Given the large dimension of the parameter space ($2^6 - 1$) a reduced

Table 4 NMC data. Selected model. MLE (ASE) of the log-hybrid linear parameters of interest, related focal coefficients and corresponding modifiers (with respect to the baseline model)

D	$\varphi_{cvd,pac,\bar{Z}_D}$	$\beta^{*,D}$	$\beta^{*,D} - \beta^*$
$\emptyset$	−0.019 (0.148)	−0.019 (0.148)	0
$\{bmi\}$	0.349 (0.207)	0.330 (0.254)	0.349 (0.207)
$\{sle\}$	0.277 (0.263)	0.258 (0.301)	0.277 (0.263)
$\{bmi, sle\}$	0	0.607 (0.361)	0.626 (0.330)

log-hybrid linear model is selected using a backward stepwise procedure where non-significant parameters are iteratively constrained to zero. The selected model includes 33 non-zero log-hybrid linear parameters and provides a satisfactory fit with a value of the log-likelihood ratio test equal to 36.28 on 30 degrees of freedom (p-value=0.20). Although in this reduced model some parameter estimates provide a p-value greater than 0.05, it has been verified that more parsimonious models with additional zero constraints are not adequate in terms of likelihood ratio test.

Table 4 collects the estimates of the log-hybrid linear parameters $\varphi_{cvd,pac,\bar{Z}_D}$ required to obtain the estimates of the logistic regression coefficients $\beta^{*,D}$, with $D \in \{bmi, sle\}$ and $*$ standing for $\{sex, age\}$. A simplified notation is used as in Sect. 6.3 such that $\varphi_{cvd,pac,\bar{Z}_D}$ stands for $\varphi^{cvd,pac,*}_{cvd,pac,\bar{Z}_D}$, $\beta^{*,D}$ stands for $\beta^{*,D}_{pac}$ and D is omitted when equal to $\emptyset$. The estimates in the third column of Table 4 represent a measure of the overall effect β^* of the treatment and of its three direct effects $\beta^{*,bmi}$, $\beta^{*,sle}$, and $\beta^{*,bmi,sle}$, which partially or fully constrain bmi and sle to their reference levels. The estimate of β^* suggests that the overall effect of the physical activity is non-significant to explain the probability of adverse cardiovascular events. Direct effects are estimated to be positive across models, supporting the hypothesis that a low physical activity increases the probability of adverse event when high body mass index and/or sleeping problems occur. The highest estimated direct effect is $\beta^{*,bmi,sle}$. By using Theorem 3, this parameter results as a linear combination of the overall effect and of the direct effects in nested models,

$$\beta^{*,bmi,sle} = -\beta^* + \beta^{*,bmi} + \beta^{*,sle}, \tag{29}$$

because $\varphi_{cvd,pac,\overline{bmi},\overline{sle}} = 0$ in the selected model.

From the application of Theorem 3, which handles the case where the base model not only includes the focal variable, the differences between the various indirect effects and the overall effect are obtained as

$$\beta^{*,D} - \beta^* = \sum_{D' \subseteq D : D' \neq \emptyset} \varphi_{cvd,pac,\bar{Z}_D}, \qquad D \subseteq \{bmi, sle\}.$$

These differences are modifiers of the focal parameters, as discussed in Sect. 6.3, and account, in the present mediation context, for the indirect effects of the treatment on the disease. Under the selected model, these differences reduce to

$$\beta^{*,bmi} - \beta^{*} = \varphi_{cvd,pac,\overline{bmi}}, \tag{30}$$

$$\beta^{*,sle} - \beta^{*} = \varphi_{cvd,pac,\overline{sle}}, \tag{31}$$

$$\beta^{*,bmi,sle} - \beta^{*} = \varphi_{cvd,pac,\overline{bmi}} + \varphi_{cvd,pac,\overline{sle}} \tag{32}$$

and the corresponding estimates are collected in the last column of Table 4. Since $\hat{\beta}^{*}$ is close to zero, the estimates of these differences approximately equal the direct effect estimates. However, as discussed in Raggi et al. (2023), it would be incautious to conclude that the direct and indirect effects have a comparable magnitude in this specific application, because the differences in (30)-(32) may embed residual effects due to the non-collapsibility of the odds ratio.

Appendix A Parameterizations of interest

The probability distribution of a binary random vector is naturally parameterized by a cross-classified probability table, that is, by the collection of cell probabilities. However, submodels of interest are identified by nonlinear, multiplicative, constraints on the probabilities and, therefore, alternative parameterizations have been introduced in such a way that relevant submodels are characterized by the vanishing of certain parameters. Here, we review the main features of three such parameterizations, as required for this paper, with special attention to their connection with the odds ratio as well as to the relationships linking one parameterization with the other. Specifically, we present log-linear parameters, logistic regression coefficients, and log-hybrid linear parameters, in three distinct subsections, after a brief common introduction.

Because of the different roles played by variables in the three parameterizations we consider, it is convenient, in this section, to use a slightly different notation from the rest of the paper, and we represent all the variables in a single random vector $Y_V = (Y_1, Y_2, \dots, Y_v)$ indexed by the set $V = \{1, 2, \dots, v\}$. We remark, however, that the notation used in the rest of the paper can immediately be recovered by setting $Y = Y_1$, $X = Y_2$ and $Z = (Y_3, Y_4 \dots, Y_v)$. Without loss of generality, we assume that the levels of the variables are labelled by 0 and 1, so that the probability distribution of Y_V is characterized by the collection of probabilities $\mathrm{pr}(Y_V = y_V)$, $y_V \in \{0, 1\}^{|V|}$, which can be equivalently written as

$$\mathrm{pr}(Y_D = 1, Y_{V\setminus D} = 0), \qquad D \subseteq V, \tag{A1}$$

where $Y_D = 1$ means $Y_i = 1$ for every $i \in D$, and similarly for $Y_{V\setminus D} = 0$. Note that each $D \subseteq V$ can be identified with the set of variables $\{Y_i\}_{i\in D}$, whenever this proves useful.

Möbius inversion formula is a well-known combinatorial tool that we extensively exploit in the derivation of results. We formally state it in the following proposition, and we refer to Roverato (2017) for a proof and more details on its application to the theory of categorical data.

Proposition A.1 *(Möbius inversion formula) Let* $\theta = (\theta_D)_{D \subseteq V}$ *and* $\omega = (\omega_D)_{D \subseteq V}$ *be two real vectors indexed by the subsets of a finite set* V*. Then, it holds that*

$$\omega_D = \sum_{D' \subseteq D} \theta_{D'} \quad \textit{if and only if} \quad \theta_D = \sum_{D' \subseteq D} (-1)^{|D \setminus D'|}\, \omega_{D'},$$

where the identities are intended for all $D \subseteq V$*.*

A.1 Log-linear parameters

In data analysis settings where all the variables are on an equal footing, it is common to apply log-linear models, which rely on a parameterization obtained from a log-linear expansion of the cell probabilities. We denote the log-linear parameters by λ, and recall that there exist different ways, although substantially equivalent, to compute the parameter values.

We follow Chapter 9 of Agresti (2012) and apply the corner point constraint with zero reference level so as to obtain

$$\log \mathrm{pr}(Y_D = 1, Y_{V \setminus D} = 0) \;=\; \sum_{D' \subseteq D} \lambda_{D'}, \qquad \text{for every } D \subseteq V.$$

Hence, the application of Möbius inversion gives immediately the formula for the computation of log-linear parameters from cell probabilities:

$$\lambda_D \;=\; \sum_{D' \subseteq D} (-1)^{|D \setminus D'|} \,\log \mathrm{pr}(Y_{D'} = 1, Y_{V \setminus D'} = 0), \qquad \text{for every } D \subseteq V. \quad \text{(A2)}$$

In turn, letting $D = \{i, j\}$ with $i \neq j$, the latter identity can be used to show that

$$\lambda_{\{i,j\}} \;=\; \log \mathrm{OR}_{Y_i Y_j | Y_{V \setminus \{i,j\}} = 0}\,. \quad \text{(A3)}$$

A.2 Logistic regression coefficients

Logistic regression is a method widely applied when the analysis requires to regress a binary variable, say Y_1, on a vector of explanatory variables, say $Y_{V \setminus \{1\}}$. It relies on expressing the logit

$$\mathrm{logit}\,\mathrm{pr}(Y_1 = 1 \mid Y_{V \setminus \{1\}} = y_{V \setminus \{1\}}) \;=\; \log \frac{\mathrm{pr}(Y_1 = 1, Y_{V \setminus \{1\}} = y_{V \setminus \{1\}})}{\mathrm{pr}(Y_1 = 0, Y_{V \setminus \{1\}} = y_{V \setminus \{1\}})} \quad \text{(A4)}$$

as the following function of the values $y_{V \setminus \{1\}}$ taken by the explanatory variables:

$$\mathrm{logit}\,\mathrm{pr}(Y_1 = 1 \mid Y_{V \setminus \{1\}} = y_{V \setminus \{1\}}) \;=\; \sum_{E \subseteq V \setminus \{1\}} \beta_E \prod_{i \in E} y_i\,.$$

We refer to Chapter 5 of Agresti (2012) for more information. Since our variables can only take the values zero and one, the above displayed equation can be written in the form

$$\operatorname{logit} \operatorname{pr}(Y_1 = 1 \mid Y_E = 1, Y_{V\setminus(E\cup\{1\})} = 0) \;=\; \sum_{E'\subseteq E} \beta_{E'}, \qquad \text{for every } E \subseteq V\setminus\{1\},$$

and, in this way, we can readily apply Möbius inversion to show that the logistic regression coefficients can be computed as

$$\beta_E \;=\; \sum_{E'\subseteq E} (-1)^{|E\setminus E'|} \operatorname{logit} \operatorname{pr}(Y_1 = 1 \mid Y_{E'} = 1, Y_{V\setminus(E'\cup\{1\})} = 0), \tag{A5}$$

for every $E \subseteq V \setminus \{1\}$. As shown, for instance, by Section 9.5 of Agresti (2012), there exists a close relationship between log-linear parameters and logistic regression coefficients, which we can formalize in our notation as

$$\beta_E \;=\; \sum_{E'\subseteq E} (-1)^{|E\setminus E'|} \log \frac{\operatorname{pr}(Y_1 = 1, Y_{E'} = 1, Y_{V\setminus(E'\cup\{1\})} = 0)}{\operatorname{pr}(Y_1 = 0, Y_{E'} = 1, Y_{V\setminus(E'\cup\{1\})} = 0)} \tag{A6}$$

$$= \sum_{D\subseteq E\cup\{1\}} (-1)^{|(E\cup\{1\})\setminus D|} \log \operatorname{pr}(Y_D = 1, Y_{V\setminus D} = 0) \;=\; \lambda_{E\cup\{1\}}, \tag{A7}$$

for every $E \subseteq V \setminus \{1\}$, where the equality in (A6) follows from (A5) by applying (A4), the first equality in (A7) follows from the fact that, for every $E' \subseteq E$, the two terms of the sum in (A7) indexed by $D = E'$ and $D = E' \cup \{1\}$ add app to the term of the sum in (A6) indexed by E', and the second equality in (A7) follows from (A2). In particular, letting $E = \{j\}$ with $j \neq 1$, and using (A3), we find

$$\beta_{\{j\}} \;=\; \log \mathrm{OR}_{Y_1 Y_j \mid Y_{V\setminus\{1,j\}}=0}\,. \tag{A8}$$

A.3 Log-hybrid linear parameters

The log-hybrid linear parameterization is more recent: it was first introduced by La Rocca and Roverato (2019); see also Roverato (2017). It depends on a predefined partition of the variables into two disjoint groups indexed by the subsets T and U of V, that is, formally, on writing $V = T \cup U$ with $T \cap U = \emptyset$. Informally, this parameterization relies on the idea that log-hybrid linear parameters are computed conditionally with respect to the variables in Y_U and marginally with respect to variables in Y_T, in a sense to be clarified below. In this way, the log-hybrid linear parameterization extends the log-linear parameterization and, indeed, coincides with the latter in the case where $U = V$, so that $T = \emptyset$. On the other hand, if $U = \emptyset$ and $T = V$, then the log-hybrid linear parameterization coincides with the log-mean linear parameterization introduced by Roverato et al. (2013).

Two steps are required to compute the log-hybrid linear parameters from the cell probabilities. The first step is the construction of a collection of probabilities, indexed by the subsets of V, computed from the distribution of specific subvectors of Y_V. Concretely, every $D \subseteq V$ uniquely identifies the subvector $Y_{U\cup D} = Y_{U\cup(D\cap T)}$ of Y_V or, in other words, the distribution of Y_V marginalized over the variables $Y_{T\setminus D}$. Then, every $D \subseteq V$ can be associated with the probability of the event $\{Y_D = 1, Y_{U\setminus D} = 0\}$, thus obtaining the collection of probabilities

$$\mathrm{pr}(Y_D = 1, Y_{U\setminus D} = 0), \qquad D \subseteq V. \tag{A9}$$

It is shown in La Rocca and Roverato (2019) that (A9) is a one-to-one transformation of the collection of cell probabilities in (A1), and therefore a valid parameterization of the distribution of Y_V. Specifically, there is a linear transformation between (A1) and (A9), and it is straightforward to compute one from the other; we refer to Chapter 6 of Roverato (2017) for details.

The second step amounts to computing the log-hybrid parameters from (A9) as

$$\varphi_D^U = \sum_{D'\subseteq D} (-1)^{|D\setminus D'|} \log \mathrm{pr}(Y_{D'} = 1, Y_{U\setminus D'} = 0), \qquad \text{for every } D \subseteq V, \tag{A10}$$

and thus, also in this case, we can apply Möbius formula to obtain the inverse transformation

$$\log \mathrm{pr}(Y_D = 1, Y_{U\setminus D} = 0) = \sum_{D'\subseteq D} \varphi_{D'}^U, \qquad \text{for every } D \subseteq V,$$

which we call the *log-hybrid linear expansion* for Y_D with respect to Y_U.

Remarkably, the computation of φ_D^U in (A10) is based on the marginal distribution of $Y_{U\cup D}$ and, therefore, it satisfies upward compatibility: its value does not change if more variables are included in the analysis, or if some variables are removed and Y_V is replaced by a subvector Y_S such that $U \cup D \subseteq S$. In particular, when $D \subseteq U$ the application of (A10) amounts to computing the log-linear parameters in (A2) on the marginal distribution of Y_U, so that (A3) gives rise to

$$\varphi_{\{i,j\}}^U = \log \mathrm{OR}_{Y_i Y_j | Y_{U\setminus\{i,j\}}=0} \qquad \text{if } \{i, j\} \subseteq U; \tag{A11}$$

see also Proposition 8.7.5 in La Rocca and Roverato (2019). Furthermore, for every $W \subseteq V \setminus (U \cup D)$, it follows directly from (A10) and the fact that every non-empty subset has the same number of even and odd subsets, that

$$\varphi_D^{U|Y_W=0} = \varphi_D^{U\cup W} \qquad \text{if } D \neq \emptyset, \tag{A12}$$

where $\varphi_D^{U|Y_W=0}$ is the log-hybrid linear parameter φ_D^U computed from the conditional probability distribution of $Y_{U\cup D}$ given $Y_W = 0$ (its marginal distribution if $W = \emptyset$).

Appendix B Proofs

B.1 Proof of Lemma 1

The result is trivially true when $D = \emptyset$ and thus in the following we assume that $D \neq \emptyset$. Then, we note that, when computing $\varphi_{XY}^{XY} = \varphi_{\{X,Y\}}^{\{X,Y\}}$, formula (A10) can equivalently be written as

$$\varphi_{XY}^{XY} = \sum_{A \subseteq \{X,Y\}} (-1)^{|\{X,Y\} \setminus A|} \log \operatorname{pr}(X = 1_{\{X \in A\}}, Y = 1_{\{Y \in A\}}),$$

where $1_{\{X \in A\}}$ is the indicator function that takes value 1 if $X \in A$ and 0 otherwise, and similarly for $1_{\{Y \in A\}}$. Accordingly, the parameter $\varphi_{XY}^{XY|Z_D=1}$ is computed in the same way, but with respect to the conditional distribution of X and Y given $Z_D = 1$, that is, we have

$$\varphi_{XY}^{XY|Z_D=1} = \sum_{A \subseteq \{X,Y\}} (-1)^{|\{X,Y\} \setminus A|} \log \operatorname{pr}(X = 1_{\{X \in A\}}, Y = 1_{\{Y \in A\}} \mid Z_D = 1).$$

We will use this equality at the end of the proof.

We now consider an arbitrary non-empty subset $D \subseteq M$ and, for every $A \subseteq \{X, Y\}$, we let $\varphi_A^{XY[D]} = \sum_{D' \subseteq D} \varphi_{A \cup D'}^{XY}$, where we identify D with $\{Z_i\}_{i \in D}$ and D' with $\{Z_i\}_{i \in D'}$. We have to show that $\varphi_{XY}^{XY[D]} = \varphi_{XY}^{XY|Z_D=1}$. If we write the log-hybrid linear expansion for (Y, X, Z_D) with respect to (X, Y), we obtain

$$\begin{aligned} \log \operatorname{pr}(X = 1_{\{X \in A\}}, Y = 1_{\{Y \in A\}}, Z_D = 1) &= \sum_{F \subseteq \{X,Y\} \cup D} \varphi_F^{XY} \\ &= \sum_{A \subseteq \{X,Y\}} \sum_{D' \subseteq D} \varphi_{A \cup D'}^{XY} \\ &= \sum_{A \subseteq \{X,Y\}} \varphi_A^{XY[D]}, \end{aligned}$$

for every $A \subseteq \{X, Y\}$. Hence, we can apply the Möbius inversion formula from Proposition A.1 to obtain the desired result:

$$\begin{aligned} \varphi_{XY}^{XY[D]} &= \sum_{A \subseteq \{X,Y\}} (-1)^{|\{X,Y\} \setminus A|} \log \operatorname{pr}(X = 1_{\{X \in A\}}, Y = 1_{\{Y \in A\}}, Z_D = 1) && \text{(B1)} \\ &= \log \frac{\operatorname{pr}(X = 1, Y = 1, Z_D = 1) \operatorname{pr}(X = 0, Y = 0, Z_D = 1)}{\operatorname{pr}(X = 1, Y = 0, Z_D = 1) \operatorname{pr}(X = 0, Y = 1, Z_D = 1)} && \text{(B2)} \\ &= \log \frac{\operatorname{pr}(X = 1, Y = 1 \mid Z_D = 1) \operatorname{pr}(X = 0, Y = 0 \mid Z_D = 1)}{\operatorname{pr}(X = 1, Y = 0 \mid Z_D = 1) \operatorname{pr}(X = 0, Y = 1 \mid Z_D = 1)} && \text{(B3)} \\ &= \sum_{A \subseteq \{X,Y\}} (-1)^{|\{X,Y\} \setminus A|} \log \operatorname{pr}(X = 1_{\{X \in A\}}, Y = 1_{\{Y \in A\}} \mid Z_D = 1), && \text{(B4)} \end{aligned}$$

where in (B2) we have explicitly written the four terms of the sum (B1), while (B3) is obtained by dividing twice both the numerator and the denominator in (B2) by $\mathrm{pr}(Z_D = 1)$, and the step from (B3) to (B4) can be seen as the inverse of the step from (B1) to (B2); it follows from (B4) that $\varphi_{XY}^{XY[D]} = \varphi_{XY}^{XY|Z_D=1}$ as required.

B.2 Proof of Theorem 1

The first equality in (18), that is $\beta_X^D = \log \mathrm{OR}_{XY|Z_D=0}$, is a well-established result, which is shown for instance in (A8) of Appendix A. We prove below the second equality.

Equation (A11) implies $\varphi_{XY}^{XY} = \log \mathrm{OR}_{XY}$ and, therefore, we have

$$\varphi_{XY}^{XY|Z_D=z_D} = \log \mathrm{OR}_{XY|Z_D=z_D} \qquad \text{for every } z_D \in \{0, 1\}^{|D|}. \tag{B5}$$

Then, it follows immediately from Lemma 1 that $\sum_{D' \subseteq D} \varphi_{XY\bar{Z}_{D'}}^{XY} = \varphi_{XY}^{XY|\bar{Z}_D=1}$, which is clearly equivalent to $\sum_{D' \subseteq D} \varphi_{XY\bar{Z}_{D'}}^{XY} = \varphi_{XY}^{XY|Z_D=0}$. The latter equality, together with (B5), implies $\sum_{D' \subseteq D} \varphi_{XY\bar{Z}_{D'}}^{XY} = \log \mathrm{OR}_{XY|Z_D=0}$, which is the second equality in (18).

B.3 Proof of Theorem 2

We apply Möbius inversion formula (Proposition A.1) to the statement of Theorem 1, written as

$$\log \mathrm{OR}_{XY|Z_D=0} = \sum_{D' \subseteq D} \varphi_{XY\bar{Z}_D}^{XY} \quad \text{for every } D \subseteq M,$$

so as to obtain, for every $D \subseteq M$,

$$\begin{aligned}
\varphi_{XY\bar{Z}_D}^{XY} &= \sum_{D' \subseteq D} (-1)^{|D \setminus D'|} \log \mathrm{OR}_{XY|Z_{D'}=0} \\
&= \sum_{D' \subset D} (-1)^{|D \setminus D'|} \log \mathrm{OR}_{XY|Z_{D'}=0} + \log \mathrm{OR}_{XY|Z_D=0}
\end{aligned}$$

and, therefore,

$$\begin{aligned}
\log \mathrm{OR}_{XY|Z_D=0} &= -\sum_{D' \subset D} (-1)^{|D \setminus D'|} \log \mathrm{OR}_{XY|Z_{D'}=0} + \varphi_{XY\bar{Z}_D}^{XY} \\
&= \sum_{D' \subset D} (-1)^{|D \setminus D'|-1} \log \mathrm{OR}_{XY|Z_{D'}=0} + \varphi_{XY\bar{Z}_D}^{XY},
\end{aligned}$$

which establishes (20).

We deduce (19) from (20) by replacing the log odds ratios with logistic regression coefficients, according to the well-known identity (A8) of Appendix A.

B.4 Proof of Theorem 3

The equalities $\beta_X^C = \log \mathrm{OR}_{XY|Z_C=0} = \varphi_{XY}^{XYZ_C}$ follow from (A8) and (A11) in Appendix A. Then, for all $C \subseteq M$, and $D \subseteq M \setminus C$, we think of $\log \mathrm{OR}_{XY|Z_C=0, Z_D=0}$ as $\log \mathrm{OR}_{XY|Z_{D\setminus C}=0}$ computed from the conditional probability distribution of X, Y and Z_D given $Z_C = 0$. This enables us to express $\log \mathrm{OR}_{XY|Z_D=0}$ as $\sum_{D' \subseteq D} \varphi_{XY\bar{Z}_{D'}}^{XY|Z_C=0}$, which equals $\sum_{D' \subseteq D} \varphi_{XY\bar{Z}_{D'}}^{XYZ_C}$ by (A12). We obtain (21) and (22) by recalling the initial equalities.

Acknowledgements Financial support was provided by the European Union - Next Generation EU, Mission 4 Component 2 - CUP B53D23009470006 and CUP C53D23002580006 (MUR-PRIN grant 2022SMNNKY). We are thankful to two anonymous Reviewers for their helpful comments, which resulted in a significantly improved paper.

Funding Open access funding provided by Università degli Studi di Modena e Reggio Emilia within the CRUI-CARE Agreement.

Declarations

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

References

Agresti A (2012) Categorical data analysis, 3rd edn. Wiley series in probability and statistics. Wiley, Hoboken

Altham PM (1970) The measurement of association of rows and columns for an r × s contingency table. J Roy Stat Soc: Ser B (Methodol) 32(1):63–73

Asmussen S, Edwards D (1983) Collapsibility and response variables in contingency tables. Biometrika 70(3):567–578

Clogg CC, Petkova E, Haritou A (1995) Statistical methods for comparing regression coefficients between models. Am J Sociol 100(5):1261–1293

Clogg CC, Petkova E, Shihadeh ES (1992) Statistical methods for analyzing collapsibility in regression models. J Educ Stat 17(1):51–74

Cox DR, Wermuth N (1996) Multivariate dependencies: models, analysis, and interpretation. Chapman & Hall, Boca Raton

Cox DR, Wermuth N (2003) A general condition for avoiding effect reversal after marginalization. J R Stat Soc Ser B Stat Methodol 65(4):937–941

Didelez V, Kreiner S, Keiding N (2010) Graphical models for inference under outcome-dependent sampling. Stat Sci 25(3):368–387

Drton M (2009) Discrete chain graph models. Bernoulli 15(3):736–753

Ducharme GR, Lepage Y (1986) Testing collapsibility in contingency tables. J R Stat Soc Ser B (Methodol) 48(2):197–205

Edwards AW (1963) The measure of association in a 2×2 table. J R Stat Soc Ser A Stat Soc 126(1):109–114

Efron B (1984) Comparing non-nested linear models. J Am Stat Assoc 79(388):791–803

Greenland S, Mickey RM (1988) Closed form and dually consistent methods for inference on strict collapsibility in $2 \times 2 \times K$ and $2 \times J \times K$ tables. J R Stat Soc Ser C (Appl Stat) 37(3):335–343

Greenland S, Pearl J, Robins JM (1999) Confounding and collapsibility in causal inference. Stat Sci 14(1):29–46

Guo J, Geng Z (1995) Collapsibility of logistic regression coefficients. J R Stat Soc Ser B (Methodol) 57(1):263–267

La Rocca L, Roverato A (2019) Discrete graphical models. In: Lauritzen SL, Maathuis M, Wainwright M (eds) Drton M. Handbook of graphical models. Handbooks of modern statistical methods. Chapman and Hall/CRC, Boca Raton

Lang JB (1996) Maximum likelihood methods for a generalized class of log-linear models. Ann Stat 24(2):726–752

Pearl J (2001) Direct and indirect effects. In: Proceedings of the seventeenth conference on uncertainty in artificial intelligence, pp. 411–420. Morgan Kaufmann, San Francisco, CA

Pearl J (2012) The causal mediation formula-a guide to the assessment of pathways and mechanisms. Prev Sci 13(4):426–436

R Core Team (2023) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Raggi M, Stanghellini E, Doretti M (2023) Path analysis for binary random variables. Sociol Methods Res 52(4):1883–1915

Roverato A (2017) Graphical models for categorical data. Cambridge University Press, SemStat Elements

Roverato A, Lupparelli M, La Rocca L (2013) Log-mean linear models for binary data. Biometrika 100(2):485–494

Royston P, Thompson SG (1995) Comparing non-nested regression models. Biometrics 51(1):114–127

Schwartz S, Li F, Mealli F (2011) A Bayesian semiparametric approach to intermediate variables in causal inference. J Am Stat Assoc 106(496):1331–1344

Sjölander A, Humphreys K, Vansteelandt S, Bellocco R, Palmgren J (2009) Sensitivity analysis for principal stratum direct effects, with an application to a study of physical activity and coronary heart disease. Biometrics 65(2):514–520

Stanghellini E, Doretti M (2019) On marginal and conditional parameters in logistic regression models. Biometrika 106(3):732–739

VanderWeele TJ (2015) Explanation in causal inference: methods for mediation and interaction. Oxford University Press, New York

Vuong QH (1989) Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57(2):307–333

Wang K (2021) Relating parameters in conditional, marginalized, and marginal logistic models when the mediator is binary. Stat Interface 14(2):109–114

Wermuth N (1987) Parametric collapsibility and the lack of moderating effects in contingency tables with a dichotomous response variable. J R Stat Soc Ser B (Methodol) 49(3):353–364

Wermuth N, Cox DR (2008) Distortion of effects caused by indirect confounding. Biometrika 95(1):17–33

Whittemore AS (1978) Collapsibility of multidimensional contingency tables. J Roy Stat Soc Ser B (Methodol) 40(3):328–340

Xie X, Ma Z, Geng Z (2008) Some association measures and their collapsibility. Stat Sin 18(3):1165–1183

Yin J, Markes S, Richardson TS, Wang L (2022) Multiplicative effect modelling: the general case. Biometrika 109(2):559–566