Journal of Multivariate Analysis 196 (2023) 105179

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

On the exploration of regression dependence structures in multidimensional contingency tables with ordinal response variables

Zheng Wei [a], Li Wang [b], Shu-Min Liao [c], Daeyoung Kim [d,*]

[a] Department of Mathematics and Statistics, Texas A&M University-Corpus Christi, Corpus Christi, TX 78412, USA
[b] Applied Scientist II, Microsoft, 14820 36th St, Redmond, WA 98052, USA
[c] Department of Mathematics and Statistics, Amherst College, Amherst, MA 01002, USA
[d] Department of Mathematics and Statistics, University of Massachusetts-Amherst, Amherst, MA 01003-9305, USA

ARTICLE INFO

Article history:
Received 15 March 2022
Received in revised form 14 March 2023
Accepted 14 March 2023
Available online 16 March 2023

AMS 2010 subject classifications:
primary 62H20
secondary 62H17

Keywords:
Conditional association
Marginal association
Multivariate contingency tables
Ordinal response
Regression

ABSTRACT

In this paper, we propose a new data-driven method to explore complex regression dependence structures in a multi-dimensional contingency table with an ordinal response variable and categorical (ordinal or nominal) explanatory variables. The proposed method is based on a sequential decomposition of the overall regression dependence for the data quantified by the checkerboard copula regression association measure (Wei and Kim, 2021) in an informative and interpretable fashion. It can measure the marginal and conditional contributions of any subset of all available explanatory variables to the overall regression association in a hierarchical manner taking into account the order of the explanatory variables. Thus, the proposed method enables a holistic description of various aspects of regression association in a multivariate contingency table, including marginal and conditional associations between an ordinal response variable and a subset of explanatory variables of interest. We investigate theoretical properties of the proposed decomposition method, and we further illustrate its performance through simulation and two real data examples, one from a randomized controlled trial and the other from a longitudinal epidemiological study.

1. Introduction

Multivariate categorical data with ordinal response variables is very common in various fields of science such as social, behavioral, health, and biomedical sciences and it is typically summarized into a multi-dimensional contingency table. A major goal of statistical analysis of such multivariate data is to uncover, model and understand regression dependence between an ordinal response variable and a set of independent categorical variables. To this end, two types of methods have been developed [2,21,30,38,50,53]. The first one is a model-based method that requires an explicit specification of the underlying regression dependence structure and is useful for the formal modeling such as explanatory modeling or predictive modeling. The second one is a non-model-based method for descriptive/exploratory analysis which does not need parametric specification of the dependence structure in the data.

The model-based method includes (but not limited to) cumulative link models [4,18,41], continuation ratio logit model [36], adjacent-categories logit model [26,48], stereotype model [5], latent variable models, and association

* Corresponding author.
E-mail address: dkim1@umass.edu (D. Kim).

https://doi.org/10.1016/j.jmva.2023.105179

models [25,27]. The non-model-based method includes Cochran–Mantel–Haenszel methods [37,40], ordinal odds ratios [15,25,26,41], rank-based methods including Kendall's tau [32], Goodman and Kruskal's gamma [28], Spearman's rank-based correlation [2], and their extensions (for example, Kendall's conditional tau-b [1,33], Goodman and Kruskal's conditional gamma [10], conditional Spearman's rank correlation using probability-scale residuals [39]), and copula-based measures such as checkerboard copula-based concordance/rank correlation measures [12,43] and the checkerboard copula regression association measure [54].

An essential step for proper analysis of multivariate categorical data with an ordinal response variable is to explore various aspects of regression dependence structures between an ordinal response variable and a set of categorical independent variables in a model-free manner. The information obtained from this step would be informative in gaining insight into the complex regression dependences for descriptive modeling, formulating patterns of regression associations or identifying potential independent variables that may provide relevant and non-redundant information on an ordinal response variable [14,47,52].

However, it is not clear how to use the non-model-based approaches listed above in a holistic and understandable way. This is because each of them was developed to measure only one of regression associations: either global association between an ordinal response variable and all available independent variables, or bivariate marginal associations, or bivariate conditional associations conditional on the remaining independent variables. It would be desirable to develop a systematic method to accomplish two goals: (1) measuring all aforementioned associations (i.e., global/marginal/conditional associations) in an integrated manner, and (2) quantifying the contributions of any subset of independent variables of interest (via marginal/conditional associations) to the measured global regression association in an interpretable manner.

To achieve these goals, we propose a new data-driven and model-free approach to the analysis of a multidimensional contingency table with an ordinal response variable and categorical (ordinal or nominal) independent variables. The main idea of the proposed approach is based on the decomposition of the overall Checkerboard Copula Regression Association Measure (CCRAM) [54] in a holistic and interpretable fashion. The overall CCRAM was developed to quantify the global regression association for the data using the checkerboard copula regression of an ordinal response variable on all available categorical independent variables. However, as claimed by Kendall and Stuart [34], the regression dependence resulting from the joint behavior of the variables is too complex to be understood in a single coefficient, and additional steps are necessary to distill meaningful information from such a global dependence. Thus, we first define the marginal and conditional CCRAMs between an ordinal response variable and a subset of independent categorical variables of interest, one unconditional and the other conditional on other independent variables. More specifically, the conditional CCRAM is designed to quantify the conditional contribution of additionally considered categorical independent variables given other independent variables are taken into account, unlike the marginal CCRAM. We then propose a sequential decomposition method for the overall CCRAM and show that the overall CCRAM can be partitioned into a sum of a marginal CCRAM and a sequence of conditional CCRAMs in a hierarchical manner which considers the order of the categorical independent variables of interest. An advantage of the proposed decomposition is that it enables us to quantify the marginal and conditional contributions of any subset of categorical independent variables of interest to the overall regression association (measured by the overall CCRAM).

The rest of the article is organized as follows. Section 2 gives a brief review on the overall CCRAM [54] for the global regression association between an ordinal response variable and all categorical (nominal/ordinal) independent variables in a multi-dimensional contingency table. In Section 3 we define the marginal and conditional CCRAMs between an ordinal response variable and a subset of independent categorical variables of interest, one unconditional and the other conditional on other independent variables, and investigate their theoretical properties. Then we propose a sequential decomposition approach of the overall CCRAM in a hierarchical manner, taking into account the order of the categorical independent variables of interest, and investigate its properties under different types of independence for a multi-way contingency table. Note that the proposed marginal/conditional CCRAMs and decomposition method are also applicable for the contingency tables with nominal independent variables. Section 4 discusses the estimation of the marginal and conditional CCRAMs and the decomposition method proposed in Section 3. Section 5 conducts data analysis to illustrate the practical utility of the proposed decomposition method with two real data examples, one from a randomized controlled trial and the other from a longitudinal epidemiological study. Section 6 closes the paper with a discussion on simulation results and future work.

2. Review on checkerboard copula regression association measure

Suppose that objects in a given population can be classified into a $(d+1)$-way contingency table with respect to $(d+1)$-dimensional ordinal vector $(\boldsymbol{X}^\top, Y)$ where Y is a response variable and $\boldsymbol{X} = (X_1, \ldots, X_k, \ldots, X_d)$ is a d-dimensional vector of independent variables. Here, Y has J ordered levels $(\{y_1 < \cdots < y_j < \cdots < y_J\})$ and X_k has I_k ordered levels $(\{x_1^k < \cdots < x_{i_k}^k < \cdots < x_{I_k}^k\})$ where $<$ represents the natural ordering of levels of each variable, $k \in \{1, \ldots, d\}$, $i_k \in \{1, \ldots, I_k\}$ and $j \in \{1, \ldots, J\}$. The joint probability mass function (p.m.f.) of $(\boldsymbol{X}^\top, Y)$ is $p_{\boldsymbol{i},j} = P(\boldsymbol{X} = \boldsymbol{x}_{\boldsymbol{i}}, Y = y_j)$, the probability that $(\boldsymbol{X}^\top, Y)$ occurs in the cell $(\boldsymbol{i}, j)$ representing a possible outcome $(\boldsymbol{x}_{\boldsymbol{i}}^\top, y_j)$ of $(\boldsymbol{X}^\top, Y)$ (or an object falls into the $(\boldsymbol{i}, j)$ cell) where $\boldsymbol{i} = (i_1, \ldots, i_k, \ldots, i_d)$ and $\boldsymbol{x}_{\boldsymbol{i}} = (x_{i_1}^1, \ldots, x_{i_k}^k, \ldots, x_{i_d}^d)$. The p.m.f. of d-dimensional marginal table associated with $\boldsymbol{X}$, denoted by $p_{\boldsymbol{i}+}$, is obtained by summing $p_{\boldsymbol{i},j}$ over all levels in Y and the p.m.f.s of 1-dimensional marginal tables for X_k and Y, denoted by p_{+i_k+} and p_{+j}, are obtained by summing $p_{\boldsymbol{i},j}$ over levels of all other factors except for X_k and Y, respectively. Then, the conditional p.m.f. for Y given $\boldsymbol{X}$ is defined to be $p_{j|\boldsymbol{i}} = p_{\boldsymbol{i},j}/p_{\boldsymbol{i}+}$. We shall denote the cumulative distribution function (c.d.f.) of $(\boldsymbol{X}^\top, Y)$ and the marginal c.d.f.s of X_k and Y to be $H_{\boldsymbol{X},Y}$, F_{X_k} and F_Y, respectively. The ranges of F_{X_k} and F_Y, denoted as D_{X_k} and D_Y, respectively, can be obtained from their respective p.m.f.s : $D_{X_k} = \{u^k_0, \ldots, u^k_{i_k}, \ldots, u^k_{I_k}\}$ and $D_Y = \{v_0, \ldots, v_j, \ldots, v_J\}$ where $u^k_{i_k} = F_{X_k}(x^k_{i_k}) = \sum_{\ell_k=1}^{i_k} p_{+\ell_k+}$, $v_j = F_Y(y_j) = \sum_{\ell=1}^{j} p_{+\ell}$, $u^k_0 = v_0 = 0$, and $u^k_{I_k} = v_J = 1$.

When the main interest of the study is to model the dependence among variables, copulas have became ubiquitous statistical tools [31,42]. By Sklar [49], there exists a function called copula C that links the marginal distributions together to form the joint distribution, i.e., $C : [0,1]^{d+1} \to [0,1]$ such that $H_{\boldsymbol{X},Y}(x_1, \ldots, x_d, y) = C(F_{X_1}(x_1), \ldots, F_{X_d}(x_d), F_Y(y))$, for all $x_1, \ldots, x_d, y \in \mathbb{R}$. When $F_{X_1}, \ldots, F_{X_d}, F_Y$ are continuous, C is unique and is the distribution function of $\boldsymbol{F}(\boldsymbol{X}, Y) = (F_{X_1}(X_1), \ldots, F_{X_d}(X_d), F_Y(Y))$. Otherwise, C is uniquely determined on the range of $\boldsymbol{F}(\boldsymbol{X}, Y)$. When X_ks and Y are ordinal variables in a contingency table, D_{X_k} and D_Y (the range of $F_{X_k}(X_k)$ and $F_Y(Y)$) are finite subsets of $[0,1]$ and the distribution of $\boldsymbol{F}(\boldsymbol{X}, Y)$ is actually a subcopula, denoted by C^S, i.e., a copula function with the domain $\left(\prod_{k=1}^d D_k\right) \times D_Y$, which is a strict subset of $[0,1]^{d+1}$. There are two main issues with the use of subcopula in characterizing the dependence among ordinal variables from the data. First, it is not an easy task to study the convergence in the space of subcopulas because along with the true subcopula, its domain also needs to be estimated from the data and its estimator varies with the sample size. This must be incorporated into a metric space characterizing the convergence of subcopula estimation [44]. Second, a subcopula cannot separate the dependence structure from the marginal distributions because its domain depends on the marginal distributions.

A common practice to overcome the issues above is to use the distributional transform of the ordinal variables [43,45]:

$$U_k = F_{X_k}(X_k-) + \{F_{X_k}(X_k) - F_{X_k}(X_k-)\}\Lambda_k, \qquad V = F_Y(Y-) + \{F_Y(Y) - F_Y(Y-)\}\Lambda_{d+1}, \tag{1}$$

where $k \in \{1, \ldots, d\}$, $F(x-)$ refers to the left limit of F, and $\Lambda_1, \ldots, \Lambda_{d+1}$ denote independent uniform random variables on $[0,1]$ independent of a (d+1)-dimensional ordinal vector $(\boldsymbol{X}^\top, Y)$. The continued random variables U_k and V in (1) are standard uniform random variables and the vector $(\boldsymbol{U}^\top, V)$ with $\boldsymbol{U} = (U_1, \ldots, U_k, \ldots, U_d)$ has a unique copula with domain $[0,1]^d$, known as the checkerboard copula, denoted as C^+ or $C^+_{\boldsymbol{U},V}$. The checkerboard copula C^+ is a smooth version of the subcopula C^S for $(\boldsymbol{X}^\top, Y)$ in a contingency table and it preserves the dependence structure from $(\boldsymbol{X}^\top, Y)$ [22,23,42,46]. Specifically, C^+ is a distribution on $[0,1]^{d+1}$ defined by subdividing the hyper-cube into small hyper-cubes $\left(\prod_{k=1}^d [u^k_{i_k-1}, u^k_{i_k}]\right) \times [v_{j-1}, v_j]$ with the corresponding constant density $c^+(\boldsymbol{u}, v)$ on each small hyper-cube [35], as defined in (2):

$$c^+(\boldsymbol{u}, v) = \frac{p_{\boldsymbol{i},j}}{\left(\prod_{k=1}^d p_{+i_k+}\right) \times p_{+j}} \quad \text{if} \quad u^k_{i_k-1} < u_k \le u^k_{i_k}, \qquad v_{j-1} < v \le v_j, \tag{2}$$

where $\boldsymbol{u} = (u_1, \ldots, u_k, \ldots, u_d) \in [0,1]^d$, $u^k_{i_k} \in D_{X_k}$ and $v_j \in D_Y$. Using the checkerboard copula density defined in (2) (as a function of the joint p.m.f. of $(\boldsymbol{X}^T, Y)$ and the marginal p.m.f.s of X_k and Y), the d-marginal density $c^+_{\boldsymbol{U}}$ for $\boldsymbol{U}$ and the conditional density $c^+_{V|\boldsymbol{U}}$ of V given $\boldsymbol{U}$ can be obtained.

Recently, Wei and Kim [54] proposed a non-model-based measure of regression association called Checkerboard Copula Regression Association Measure (CCRAM) for a multi-way contingency table with an ordinal response variable and a set of categorical (nominal/ordinal) independent variables. The CCRAM is designed to quantify the strength of overall regression dependence identified by the checkerboard copula regression of an ordinal response variable on all available categorical independent variables. In the following, we briefly review the CCRAM and its basic theoretical properties.

Definition 1 (*[54]*)**.** For the contingency table of $(\boldsymbol{X}^\top, Y)$ with the ordinal response variable Y, the checkerboard copula regression of Y on $\boldsymbol{X}$ via the continued variables $(\boldsymbol{U}^\top, V)$ in (1) is defined as follows: for $u^k_{i_k-1} < u_k \le u^k_{i_k}$ and $k \in \{1, \ldots, d\}$,

$$r_{V|\boldsymbol{U}}(\boldsymbol{u}) \equiv \mathrm{E}_{c^+_{V|\boldsymbol{U}}}(V|\boldsymbol{U} = \boldsymbol{u}) = \int_0^1 v c^+_{V|\boldsymbol{U}}(v|\boldsymbol{u}) dv = \sum_{j=1}^J s_j p_{j|\boldsymbol{i}}, \tag{3}$$

where $\{s_j\}_{j=1}^J$ with $s_j = (v_{j-1} + v_j)/2$ and $v_j = F_Y(y_j) = \sum_{\ell=1}^j p_{+\ell}$ is the checkerboard copula score set for Y, preserving the natural ordering of the categorical scale in Y.

The overall Checkerboard Copula Regression Association Measure (CCRAM) of Y on $\boldsymbol{X}$ via the continued variables $(\boldsymbol{U}^\top, V)$ in (1) is

$$\rho^2_{(\boldsymbol{X}\to Y)} \equiv \frac{\mathrm{Var}\{r_{V|\boldsymbol{U}}(\boldsymbol{U})\}}{\mathrm{Var}(V)} = \frac{\mathrm{E}\left[\left\{r_{V|\boldsymbol{U}}(\boldsymbol{U}) - 1/2\right\}^2\right]}{1/12} = 12 \sum_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{I}} \left(\sum_{j=1}^J s_j p_{j|\boldsymbol{i}} - 1/2\right)^2 p_{\boldsymbol{i}+}, \tag{4}$$

where $\mathbf{1} = (1, \dots, 1)$ is of length d and $\boldsymbol{I} = (I_1, \dots, I_d)$. The overall Scaled CCRAM (SCCRAM) of Y on $\boldsymbol{X}$ (the scaled version of the CCRAM in (4)) is

$$\rho^{2\,*}_{(\boldsymbol{X}\to Y)} = \frac{\rho^2_{(\boldsymbol{X}\to Y)}}{12\sigma_Y^2}, \tag{5}$$

where $\sigma_Y^2 = \sum_{j=1}^{J} \upsilon_{j-1}\upsilon_j/4p_{+j}$ denotes the variance of Y equipped with its checkerboard copula score, $\{s_j\}_{j=1}^{J}$.

The checkerboard copula regression $r_{V|\boldsymbol{U}}(\boldsymbol{u})$ in (3) can be viewed as the mean checkerboard score of Y with respect to its conditional distribution given a combination of categories in $\boldsymbol{X}$. The overall CCRAM $\rho^2_{(\boldsymbol{X}\to Y)}$ in (4) can quantify the linear/nonlinear relationship between Y and $\boldsymbol{X}$ detected by the checkerboard copula regression in (3). The range of $\rho^2_{(\boldsymbol{X}\to Y)}$ is between zero and an upper bound $12\sigma_Y^2$. The $\rho^2_{(\boldsymbol{X}\to Y)}$ represents the lower bound on the average proportion of variance for Y (with respect to its checkerboard copula score and its marginal distribution) accounted by the checkerboard copula regression $r_{V|\boldsymbol{U}}(\boldsymbol{u})$. A zero value of $\rho^2_{(\boldsymbol{X}\to Y)}$ implies that the mean checkerboard copula score of Y in (3) is constant over every combination of categories in $\boldsymbol{X}$, i.e., no contribution of independent variables in $\boldsymbol{X}$ to the construction of the checkerboard copula regression. $\rho^2_{(\boldsymbol{X}\to Y)}$ is equal to its upper bound $12\sigma_Y^2$ if and only if Y is a function of $\boldsymbol{X}$ almost surely. If one wishes to evaluate whether the regression association measured by $\rho^2_{(\boldsymbol{X}\to Y)}$ in (4) is strong or weak, it is necessary to use the SCCRAM ("scaled" version of the CCRAM), $\rho^{2\,*}_{(\boldsymbol{X}\to Y)}$ in (5), which ranges from 0 to 1. Note that $\rho^2_{(\boldsymbol{X}\to Y)}$ and $\rho^{2\,*}_{(\boldsymbol{X}\to Y)}$ are invariant in the permutations of categories in each X_k and/or binary Y, and thus they can also be applied when any independent variables in $\boldsymbol{X}$ are nominal and/or Y is binary.

3. Sequential decomposition of overall CCRAM and SCCRAM

In this section we propose a new decomposition approach of the overall CCRAM in (4) and the overall SCCRAM in (5) which enables us to explore various aspects of the regression association structure for a multi-dimensional contingency table with an ordinal response variable. We first define the marginal and conditional CCRAMs/SCCRAMs between an ordinal response variable and a subset of independent categorical variables of interest, one unconditional and the other conditional on other independent variables. Then we decompose the overall CCRAM/SCCRAM into a marginal CCRAM/SCCRAM and a sequence of conditional CCRAMs/SCCRAMs to quantify the marginal and conditional contributions of a set of categorical independent variables.

For the sake of a clear presentation, we start in Section 3.1 with a three-dimensional contingency table with two categorical independent variables and an ordinal dependent variable. We then generalize in Section 3.2 the derived results to multi-dimensional contingency tables with d categorical independent variables.

3.1. Decomposition for a three-way contingency table

Let X_1, X_2, and Y be three ordinal variables in a contingency table with the joint p.m.f. $p_{i_1,i_2,j}$. Then, according to Definition 1, the overall CCRAM of Y on (X_1, X_2) is given by

$$\rho^2_{(X_1,X_2\to Y)} = \frac{\mathrm{Var}\{r_{V|U_1,U_2}(U_1,U_2)\}}{\mathrm{Var}(V)} = 12\sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\left(\sum_{j=1}^{J} s_j p_{j|i_1,i_2} - 1/2\right)^2 p_{i_1,i_2,+}, \tag{6}$$

where $r_{V|U_1,U_2}(u_1,u_2) = \sum_{j=1}^{J} s_j p_{j|i_1,i_2}$ for $u^1_{i_1-1} < u_1 \le u^1_{i_1}$ and $u^2_{i_2-1} < u_2 \le u^2_{i_2}$, is the checkerboard copula regression of Y on (X_1, X_2) via the continued variables (U_1, U_2, V) in (1).

Theorem 1 presents the sequential decomposition of $\rho^2_{(X_1,X_2\to Y)}$ in (6).

Theorem 1. *Define the marginal CCRAMs of Y on X_1 and of Y on X_2, respectively:*

$$\rho^2_{(X_1\to Y)} = \frac{\mathrm{Var}\{r_{V|U_1}(U_1)\}}{\mathrm{Var}(V)} = 12\mathrm{E}\left[\left\{r_{V|U_1}(U_1) - 1/2\right\}^2\right] = 12\sum_{i_1=1}^{I_1}\left(\sum_{j=1}^{J} s_j p_{j|i_1} - 1/2\right)^2 p_{i_1++},$$

$$\rho^2_{(X_2\to Y)} = \frac{\mathrm{Var}\{r_{V|U_2}(U_2)\}}{\mathrm{Var}(V)} = 12\mathrm{E}\left[\left\{r_{V|U_2}(U_2) - 1/2\right\}^2\right] = 12\sum_{i_2=1}^{I_2}\left(\sum_{j=1}^{J} s_j p_{j|i_2} - 1/2\right)^2 p_{+i_2+},$$

where $r_{V|U_1}(u_1) = \sum_{j=1}^{J} s_j p_{j|i_1}$ for $u^1_{i_1-1} < u_1 \le u^1_{i_1}$ and $r_{V|U_2}(u_2) = \sum_{j=1}^{J} s_j p_{j|i_2}$ for $u^2_{i_2-1} < u_2 \le u^2_{i_2}$ are the checkerboard copula regression of Y on X_1 and X_2, respectively.

Define the conditional CCRAMs of Y on X_2 given X_1 and of Y on X_1 given X_2, respectively:

$$\rho^2_{(X_2\to Y|X_1)} = \frac{\mathrm{E}\left[Var\left[r_{V|U_1,U_2}(U_1,U_2)\middle|U_1\right]\right]}{\mathrm{Var}(V)} = 12\sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\left(\sum_{j=1}^{J}s_j p_{j|i_1,i_2} - \sum_{j=1}^{J}s_j p_{j|i_1}\right)^2 p_{i_1,i_2+}, \tag{7}$$

$$\rho^2_{(X_1\to Y|X_2)} = \frac{\mathrm{E}\left[Var\left[r_{V|U_1,U_2}(U_1,U_2)\middle|U_2\right]\right]}{\mathrm{Var}(V)} = 12\sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\left(\sum_{j=1}^{J}s_j p_{j|i_1,i_2} - \sum_{j=1}^{J}s_j p_{j|i_2}\right)^2 p_{i_1,i_2+}. \tag{8}$$

Then,

$$\rho^2_{(X_1,X_2\to Y)} = \rho^2_{(X_1\to Y)} + \rho^2_{(X_2\to Y|X_1)} = \rho^2_{(X_2\to Y)} + \rho^2_{(X_1\to Y|X_2)}, \tag{9}$$

and thus,

$$\rho^2_{(X_1,X_2\to Y)} \geq \rho^2_{(X_1\to Y)}, \qquad \rho^2_{(X_1,X_2\to Y)} \geq \rho^2_{(X_2\to Y)}. \tag{10}$$

The equalities in (10) hold if and only if $r_{V|U_1,U_2}(u_1,u_2) \equiv r_{V|U_1}(u_1)$ and $r_{V|U_1,U_2}(u_1,u_2) \equiv r_{V|U_2}(u_2)$ for any $u_1, u_2 \in [0,1]$, which means that the mean checkerboard score for Y with respect to the conditional distribution of Y given (X_1, X_2) is identical to that with respect to the conditional distribution of Y given X_1 (X_2) only, regardless of X_2 (X_1).

Proof. This theorem is a special case of Theorem 2, where its proof is given in Appendix A.1. □

Since $\rho^2_{(X_2\to Y|X_1)}$ in (7) and $\rho^2_{(X_1\to Y|X_2)}$ in (8) are non-negative, we see from (9) and (10) that $\rho^2_{(X_1,X_2\to Y)}$ is non-decreasing with the addition of a new variable. The proposed conditional association indices in (7) and (8) measure the contribution of a newly added variable to the global regression association measured by $\rho^2_{(X_1,X_2\to Y)}$.

We investigate the properties of the decomposition of the overall CCRAM in Theorem 1 under different types of independence available in a three-way contingency table: joint, marginal and conditional independence.

Proposition 1. *(i) If Y is jointly independent of X_1 and X_2 (i.e., $p_{i_1,i_2,j} = p_{i_1,i_2,+}p_{+j}$ for all combinations of i_1, i_2 and j), then $\rho^2_{(X_1,X_2\to Y)} = 0$ and hence the marginal and conditional association measures in (9) are all zero.*

(ii) If Y is marginally independent of X_1 in the marginal contingency table for (X_1, Y) (i.e., $p_{i_1,+,j} = p_{i_1++}p_{++j}$ for all combinations of i_1 and j), then $\rho^2_{(X_1\to Y)} = 0$ and then $\rho^2_{(X_1,X_2\to Y)} = \rho^2_{(X_2\to Y|X_1)}$. Similarly, if Y is marginally independent of X_2 in the marginal contingency table for (X_2, Y) (i.e., $p_{+,i_2,j} = p_{+i_2+}p_{+j}$ for all combinations of i_2 and j), then $\rho^2_{(X_2\to Y)} = 0$ and then $\rho^2_{(X_1,X_2\to Y)} = \rho^2_{(X_1\to Y|X_2)}$.

(iii) If Y is conditionally independent of X_1 given X_2 (i.e., $p_{j|i_1,i_2} = p_{j|i_2}$ for all combinations of i_1, i_2 and j), then $\rho^2_{(X_1\to Y|X_2)} = 0$ and $\rho^2_{(X_1,X_2\to Y)} = \rho^2_{(X_2\to Y)}$; Similarly, if Y is conditionally independent of X_2 given X_1 (i.e., $p_{j|i_1,i_2} = p_{j|i_1}$ for all combinations of i_1, i_2 and j), then $\rho^2_{(X_2\to Y|X_1)} = 0$ and $\rho^2_{(X_1,X_2\to Y)} = \rho^2_{(X_1\to Y)}$.

Proof. This theorem is a special case of Proposition 5, where its proof is given in Appendix A.2. □

Proposition 1(i) shows that, if Y is jointly independent of X_1 and X_2, the overall, conditional, and marginal association measures are all zero. Proposition 1(ii) indicates that for the marginal independence, the overall CCRAM is the same as the conditional CCRAM. Proposition 1(iii) implies that, under the conditional independence, there is no additional contribution by a newly added independent variable to the overall CCRAM.

Proposition 2 investigates the implication of zero values for the overall CCRAM and conditional CCRAM.

Proposition 2. *(i) If $\rho^2_{(X_1,X_2\to Y)} = 0$, then $r_{V|U_1,U_2}(u_1,u_2) \equiv r_{V|U_1}(u_1) \equiv r_{V|U_2}(u_2) \equiv 1/2$ for every $u_1, u_2 \in [0,1]$.*

(ii) If $\rho^2_{(X_2\to Y|X_1)} = 0$ (and thus $\rho^2_{(X_1,X_2\to Y)} = \rho^2_{(X_1\to Y)}$), then $r_{V|U_1,U_2}(u_1,u_2) = r_{V|U_1}(u_1)$ for every u_1 and u_2. Similarly, if $\rho^2_{(X_1\to Y|X_2)} = 0$ (and thus $\rho^2_{(X_1,X_2\to Y)} = \rho^2_{(X_2\to Y)}$), then $r_{V|U_1,U_2}(u_1,u_2) = r_{V|U_2}(u_2)$ for every u_1 and u_2.

Proof. Proposition 2 is a special case of Proposition 6, where its proof is given in Appendix A.3. □

According to Proposition 2(i), $\rho^2_{(X_1,X_2\to Y)} = 0$ implies that the mean checkerboard copula score in Y with respect to the conditional distribution of Y given X_1 or X_2 or both is constant over each combination of categories in X_1 and X_2. The zero value of the conditional CCRAM in Proposition 2(ii) indicates that there is no additional contribution of a newly added independent variable to the construction of the checkerboard copula regression of Y on X_1 and X_2.

3.2. Decomposition for a multi-way contingency table

Consider a $(d+1)$-way contingency table with ordinal variables $(\boldsymbol{X}^\top, Y)$ and their corresponding joint p.m.f. $p_{\boldsymbol{i},j}$ where $\boldsymbol{X} = (X_1, \ldots, X_k, \ldots, X_d)$, $\boldsymbol{i} = (i_1, \ldots, i_k, \ldots, i_d)$, $i_k \in \{1, \ldots, I_k\}$ and $j \in \{1, \ldots, J\}$. Suppose one computes the overall CCRAM, $\rho^2_{(\boldsymbol{X}\to Y)}$ in (4).

Let $\boldsymbol{X}^M_{(1)}, \ldots, \boldsymbol{X}^M_{(\ell)}, \ldots, \boldsymbol{X}^M_{(L)}$, $\ell \in \{1, \ldots, L\}$, $1 \le L \le d$, form a partition M of independent variables in $\boldsymbol{X}$ where the sub-index in $\boldsymbol{X}^M_{(\ell)}$ indicates the order of independent variables to be considered in the overall CCRAM. Denote $\boldsymbol{X}^Q_{(\ell)} = \boldsymbol{X}^M_{(1)} \bigcup \ldots \bigcup \boldsymbol{X}^M_{(\ell)}$ be the union of the first ℓ sets in the partition where $\boldsymbol{X}^Q_{(1)} = \boldsymbol{X}^M_{(1)}$ and $\boldsymbol{X}^Q_{(L)} = \boldsymbol{X}$. Accordingly, $\boldsymbol{U}^M_{(\ell)}$ and $\boldsymbol{U}^Q_{(\ell)}$ denote the continued variables of $\boldsymbol{X}^M_{(\ell)}$ and $\boldsymbol{X}^Q_{(\ell)}$ via distributional transform in (1). Let $\boldsymbol{I}^M_{(\ell)}$, $\boldsymbol{i}^M_{(\ell)}$, $\boldsymbol{I}^Q_{(\ell)}$, $\boldsymbol{i}^Q_{(\ell)}$ be index tuples for $\boldsymbol{X}^M_{(\ell)}$ and $\boldsymbol{X}^Q_{(\ell)}$ as of $\boldsymbol{I}$ and $\boldsymbol{i}$ for $\boldsymbol{X}$. Following the case of the overall CCRAM decomposition for a three-way contingency table shown in Section 3.1, we define the marginal and conditional CCRAMs below.

Definition 2. The marginal CCRAM of Y on $\boldsymbol{X}^Q_{(\ell)}$ is defined by

$$\rho^2_{(\boldsymbol{X}^Q_{(\ell)}\to Y)} = \frac{\mathrm{Var}\left\{r_{V|\boldsymbol{U}^Q_{(\ell)}}(\boldsymbol{U}^Q_{(\ell)})\right\}}{\mathrm{Var}(V)} = 12\mathrm{E}\left[\left\{r_{V|\boldsymbol{U}^Q_{(\ell)}}(\boldsymbol{U}^Q_{(\ell)}) - 1/2\right\}^2\right] = 12\sum_{\boldsymbol{i}^Q_{(\ell)}=\boldsymbol{1}}^{\boldsymbol{I}^Q_{(\ell)}}\left(\sum_{j=1}^{J} s_j p_{j|\boldsymbol{i}^Q_{(\ell)}} - 1/2\right)^2 p_{\boldsymbol{i}^Q_{(\ell)}+}, \tag{11}$$

and the conditional CCRAM of Y on $\boldsymbol{X}^M_{(\ell+1)}$ given $\boldsymbol{X}^Q_{(\ell)}$ is defined by

$$\begin{aligned}\rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} &= \frac{\mathrm{E}\left[\mathrm{Var}\left[r_{V|\boldsymbol{U}^Q_{(\ell)},\boldsymbol{U}^M_{(\ell+1)}}\left(\boldsymbol{U}^Q_{(\ell)}, \boldsymbol{U}^M_{(\ell+1)}\right)\,\middle|\,\boldsymbol{U}^Q_{(\ell)}\right]\right]}{\mathrm{Var}(V)} = \frac{\mathrm{E}\left[\left\{r_{V|\boldsymbol{U}^Q_{(\ell)},\boldsymbol{U}^M_{(\ell+1)}}(\boldsymbol{U}^Q_{(\ell)}, \boldsymbol{U}^M_{(\ell+1)}) - r_{V|\boldsymbol{U}^Q_{(\ell)}}(\boldsymbol{U}^Q_{(\ell)})\right\}^2\right]}{\mathrm{Var}(V)} \\ &= 12\sum_{\boldsymbol{i}^Q_{(\ell)}=\boldsymbol{1}}^{\boldsymbol{I}^Q_{(\ell)}}\sum_{\boldsymbol{i}^M_{(\ell+1)}=\boldsymbol{1}}^{\boldsymbol{I}^M_{(\ell+1)}}\left(\sum_{j=1}^{J} s_j p_{j|\boldsymbol{i}^Q_{(\ell)},\boldsymbol{i}^M_{(\ell+1)}} - \sum_{j=1}^{J} s_j p_{j|\boldsymbol{i}^Q_{(\ell)}}\right)^2 p_{\boldsymbol{i}^Q_{(\ell)},\boldsymbol{i}^M_{(\ell+1)}+},\end{aligned} \tag{12}$$

where $\ell \in \{1, \ldots, L-1\}$, $r_{V|\boldsymbol{U}^Q_{(\ell)}}(\boldsymbol{u}^Q_{(\ell)})$ is the checkerboard copula regression of Y on $\boldsymbol{X}^Q_{(\ell)}$ via the continued variables V and $\boldsymbol{U}^Q_{(\ell)}$ defined as

$$r_{V|\boldsymbol{U}^Q_{(\ell)}}(\boldsymbol{u}^Q_{(\ell)}) = \sum_{j=1}^{J} s_j p_{j|\boldsymbol{i}^Q_{(\ell)}} \quad \text{for} \quad u^k_{i_k-1} < u_k \le u^k_{i_k} \quad \text{with} \quad i_k \in \boldsymbol{i}^Q_{(\ell)},$$

and $r_{V|\boldsymbol{U}^Q_{(\ell)},\boldsymbol{U}^M_{(\ell+1)}}(\boldsymbol{u}^Q_{(\ell)}, \boldsymbol{u}^M_{(\ell+1)})$ is the checkerboard copula regression of Y on $(\boldsymbol{X}^Q_{(\ell)}, \boldsymbol{X}^M_{(\ell+1)})$ via the continued variables V and $(\boldsymbol{U}^Q_{(\ell)}, \boldsymbol{U}^M_{(\ell+1)})$ defined as

$$r_{V|\boldsymbol{U}^Q_{(\ell)},\boldsymbol{U}^M_{(\ell+1)}}(\boldsymbol{u}^Q_{(\ell)}, \boldsymbol{u}^M_{(\ell+1)}) = \sum_{j=1}^{J} s_j p_{j|\boldsymbol{i}^Q_{(\ell)},\boldsymbol{i}^M_{(\ell+1)}} \quad \text{for} \quad u^k_{i_k-1} < u_k \le u^k_{i_k} \quad \text{with} \quad i_k \in \boldsymbol{i}^Q_{(\ell)} \text{ or } i_k \in \boldsymbol{i}^M_{(\ell+1)}.$$

Note that the conditional CCRAM is designed to quantify the additional contribution of a set of independent variables $\boldsymbol{X}^M_{(\ell+1)}$ newly added to the checkerboard copula regression of Y on $\boldsymbol{X}^Q_{(\ell)}$.

Proposition 3 shows the monotone property of the marginal CCRAM in (11) and the additive property of the conditional CCRAM in (12).

Proposition 3.

(i) $\rho^2_{(\boldsymbol{X}^M_{(1)}\to Y)} = \rho^2_{(\boldsymbol{X}^Q_{(1)}\to Y)} \le \rho^2_{(\boldsymbol{X}^Q_{(2)}\to Y)} \le \cdots \le \rho^2_{(\boldsymbol{X}^Q_{(L-1)}\to Y)} \le \rho^2_{(\boldsymbol{X}^Q_{(L)}\to Y)} = \rho^2_{(\boldsymbol{X}\to Y)}$ *where, for* $\ell \in \{1, \ldots, L-1\}$, $\rho^2_{(\boldsymbol{X}^Q_{(\ell)}\to Y)} = \rho^2_{(\boldsymbol{X}^Q_{(\ell+1)}\to Y)}$ *if and only if* $\rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} = 0$, *i.e.,* $r_{V|\boldsymbol{U}^Q_{(\ell)},\boldsymbol{U}^M_{(\ell+1)}}(\boldsymbol{u}^Q_{(\ell)}, \boldsymbol{u}^M_{(\ell+1)}) = r_{V|\boldsymbol{U}^Q_{(\ell)}}(\boldsymbol{u}^Q_{(\ell)})$ *for any* $\boldsymbol{u}^Q_{(\ell)}$ *and* $\boldsymbol{u}^M_{(\ell+1)}$.

(ii) For each $\ell \in \{1, \ldots, L-2\}$, *let* M^* *be a new partition of* $\boldsymbol{X}$ *such that* $\boldsymbol{X}^{M^*} = \boldsymbol{X}^M_{(\ell+1)} \bigcup \boldsymbol{X}^M_{(\ell+2)}$. *Then we have*

$$\rho^2_{(\boldsymbol{X}^{M^*}\to Y|\boldsymbol{X}^Q_{(\ell)})} = \rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} + \rho^2_{(\boldsymbol{X}^M_{(\ell+2)}\to Y|\boldsymbol{X}^Q_{(\ell+1)})}, \tag{13}$$

where $\boldsymbol{X}^Q_{(\ell+1)} = \boldsymbol{X}^M_{(\ell+1)} \bigcup \boldsymbol{X}^Q_{(\ell)}$.

Proof. See Appendix A.1. □

Proposition 3(i) means that the marginal CCRAM is non-decreasing with newly added independent variables. According to Proposition 3(ii), if M^* is a refinement of a partition M such that $\boldsymbol{X}^{M^*} = \boldsymbol{X}^M_{(\ell+1)} \bigcup \boldsymbol{X}^M_{(\ell+2)}$, the conditional CCRAM of Y on $\boldsymbol{X}^{M^*}$ given $\boldsymbol{X}^Q_{(\ell)}$ is the sum of the conditional CCRAM of Y on $\boldsymbol{X}^M_{(\ell+1)}$ given $\boldsymbol{X}^Q_{(\ell)}$ and the conditional CCRAM of Y on $\boldsymbol{X}^M_{(\ell+2)}$ given $\boldsymbol{X}^Q_{(\ell+1)}$.

Proposition 4 shows that the marginal and conditional CCRAMs are invariant in the permutations of categories in each independent variable X_k and a binary response variable Y. This means that the marginal and conditional CCRAMs can also be used when any independent variables in $\boldsymbol{X}$ are nominal and/or Y is binary.

Proposition 4. *Let $\widetilde{\boldsymbol{X}}$ denote the d-dimensional categorical vector with the reordered categories for X_k from the original categories of X_k in $\boldsymbol{X}$. That is, $\widetilde{\boldsymbol{X}} = \{g_1(X_1), \ldots, g_d(X_d)\}$, where g_k is the injective function over the domains of X_k for $k \in \{1, \ldots, d\}$. If Y is a binary variable with the domain $\{y_1, y_2\}$, then we denote $\widetilde{Y}$ to be a binary variable with domain $\{\widetilde{y}_1 = y_2, \widetilde{y}_2 = y_1\}$. Then, the marginal and conditional CCRAMs are invariant with respect to $\widetilde{\boldsymbol{X}}$ and/or $\widetilde{Y}$. That is, for each $\ell \in \{1, \ldots, L-1\}$, we have*

$$\rho^2_{(\widetilde{\boldsymbol{X}}^Q_{(\ell)} \to Y)} = \rho^2_{(\boldsymbol{X}^Q_{(\ell)} \to Y)}, \quad \rho^2_{(\boldsymbol{X}^Q_{(\ell)} \to \widetilde{Y})} = \rho^2_{(\boldsymbol{X}^Q_{(\ell)} \to Y)}, \quad \rho^2_{(\widetilde{\boldsymbol{X}}^M_{(\ell+1)} \to Y | \widetilde{\boldsymbol{X}}^Q_{(\ell)})} = \rho^2_{(\boldsymbol{X}^M_{(\ell+1)} \to Y | \boldsymbol{X}^Q_{(\ell)})}, \quad \rho^2_{(\boldsymbol{X}^M_{(\ell+1)} \to \widetilde{Y} | \boldsymbol{X}^Q_{(\ell)})} = \rho^2_{(\boldsymbol{X}^M_{(\ell+1)} \to Y | \boldsymbol{X}^Q_{(\ell)})},$$

where $\widetilde{\boldsymbol{X}}^Q_{(\ell)} = \widetilde{\boldsymbol{X}}^M_{(1)} \bigcup \ldots \bigcup \widetilde{\boldsymbol{X}}^M_{(\ell)}$ and $\widetilde{\boldsymbol{X}}^M_{(1)}, \ldots, \widetilde{\boldsymbol{X}}^M_{(\ell)}, \ldots, \widetilde{\boldsymbol{X}}^M_{(L)}$ is a partition M of independent variables in $\widetilde{\boldsymbol{X}}$.

Proof. See Appendix A.1. □

Theorem 2 presents the proposed *sequential decomposition* of the overall CCRAM $\rho^2_{(\boldsymbol{X}\to Y)}$ designed to measure the contributions of any subset of independent variables considered in the computation of $\rho^2_{(\boldsymbol{X}\to Y)}$ in a hierarchical fashion taking into account the order of independent variables of interest.

Theorem 2.

(a) The sequential decomposition of the overall CCRAM $\rho^2_{(\boldsymbol{X}\to Y)}$ is given by

$$\rho^2_{(\boldsymbol{X}\to Y)} = \rho^2_{(\boldsymbol{X}^M_{(1)} \to Y)} + \sum_{\ell=1}^{L-1} \rho^2_{(\boldsymbol{X}^M_{(\ell+1)} \to Y | \boldsymbol{X}^Q_{(\ell)})}, \tag{14}$$

(b) The sequential decomposition of the overall CCRAM in (14) is invariant with respect to the permutation of the categories of any independent variable X_k in $\boldsymbol{X}$ and/or a binary dependent variable Y.

Proof. See Appendix A.1. □

As an extension of (9) for a three-way contingency table, the sequential decomposition in (14) shows that the overall CCRAM sums up one marginal CCRAM and $(L-1)$ conditional CCRAMs based on the order of the independent variables entered into the checkerboard copula regression of $\rho^2_{(\boldsymbol{X}\to Y)}$. Each of the $(L-1)$ conditional CCRAM quantifies the contribution of a newly added set of independent variables $\boldsymbol{X}^M_{(\ell+1)}$ to the overall CCRAM given all previous independent variables $\boldsymbol{X}^Q_{(\ell)}$. For example, suppose we consider a 7-dimensional contingency table with an ordinal response variable Y and six categorical independent variables $\boldsymbol{X} = (X_1, X_2, X_3, X_4, X_5, X_6)$, and assume that the order of the six independent variables considered in $\rho^2_{(\boldsymbol{X}\to Y)}$ and its partition are $\boldsymbol{X}^M_{(1)} = \{X_5, X_3\}$, $\boldsymbol{X}^M_{(2)} = \{X_1\}$, $\boldsymbol{X}^M_{(3)} = \{X_2, X_4, X_6\}$. According to (14), we can then obtain the sequential decomposition of $\rho^2_{(\boldsymbol{X}\to Y)}$ as follows:

$$\rho^2_{(\boldsymbol{X}\to Y)} = \rho^2_{(\boldsymbol{X}^M_{(1)} \to Y)} + \rho^2_{(\boldsymbol{X}^M_{(2)} \to Y | \boldsymbol{X}^Q_{(1)})} + \rho^2_{(\boldsymbol{X}^M_{(3)} \to Y | \boldsymbol{X}^Q_{(2)})} = \rho^2_{(X_5, X_3 \to Y)} + \rho^2_{(X_1 \to Y | X_5, X_3)} + \rho^2_{(X_2, X_4, X_6 \to Y | X_5, X_3, X_1)},$$

where $\boldsymbol{X}^Q_{(1)} = \boldsymbol{X}^M_{(1)} = \{X_5, X_3\}$ and $\boldsymbol{X}^Q_{(2)} = \boldsymbol{X}^M_{(1)} \cup \boldsymbol{X}^M_{(2)} = \{X_5, X_3, X_1\}$.

Note that as shown in Theorem 2(b), the sequential decomposition is valid even the independent variables are nominal or a response variable is binary.

In Proposition 5, the properties of the sequential decomposition proposed in Theorem 2(a) are investigated under different types of independence: joint, marginal and conditional independence. Note that Proposition 5 generalizes the results shown in Proposition 1.

Proposition 5.

(i) If Y is jointly independent with $\boldsymbol{X}$, then $\rho^2_{(\boldsymbol{X}\to Y)} = 0$, and hence marginal and conditional CCRAMs are all zero.

(ii) If there is a ℓ^* with $1 \le \ell^* \le L-1$ such that $\boldsymbol{X}^Q_{(\ell^*)}$ is marginally independent with Y in the marginal contingency table for $(\boldsymbol{X}^{Q\top}_{(\ell^*)}, Y)$, then $\rho^2_{(\boldsymbol{X}^M_{(1)}\to Y)} = 0$ and $\rho^2_{(\boldsymbol{X}^M_{(2)}\to Y|\boldsymbol{X}^Q_{(1)})} = \rho^2_{(\boldsymbol{X}^M_{(3)}\to Y|\boldsymbol{X}^Q_{(2)})} = \cdots = \rho^2_{(\boldsymbol{X}^M_{(\ell^*)}\to Y|\boldsymbol{X}^Q_{(\ell^*-1)})} = 0$ (note that if $\ell^*=1$, we only have $\rho^2_{(\boldsymbol{X}^M_{(1)}\to Y)} = 0$). Thus, the sequential decomposition in (14) reduces to

$$\rho^2_{(\boldsymbol{X}\to Y)} = \sum_{\ell=\ell^*}^{L-1} \rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})}.$$

(iii) If there is a ℓ^* with $1 \le \ell^* \le L-1$ such that $\boldsymbol{X}^M_{(\ell^*+1)}$ and Y are conditionally independent given $\boldsymbol{X}^Q_{(\ell^*)}$, then $\rho^2_{(\boldsymbol{X}^M_{(\ell^*+1)}\to Y|\boldsymbol{X}^Q_{(\ell^*)})} = 0$ and thus, the sequential decomposition in (14) reduces to

$$\begin{aligned}
\rho^2_{(\boldsymbol{X}\to Y)} &= \rho^2_{(\boldsymbol{X}^M_{(1)}\to Y)} + \sum_{\ell=\ell^*+1}^{L-1} \rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} && \text{for} \quad \ell^* = 1,\\
&= \rho^2_{(\boldsymbol{X}^M_{(1)}\to Y)} + \sum_{\ell=1}^{\ell^*-1} \rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} + \sum_{\ell=\ell^*+1}^{L-1} \rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} && \text{for} \quad 2 \le \ell^* \le L-2,\\
&= \rho^2_{(\boldsymbol{X}^M_{(1)}\to Y)} + \sum_{\ell=1}^{\ell^*-1} \rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} && \text{for} \quad \ell^* = L-1.
\end{aligned}$$

Proof. See Appendix A.2. □

Proposition 6 explores the property of the converse of Proposition 5.

Proposition 6.

(i) If $\rho^2_{(\boldsymbol{X}\to Y)} = 0$, then for every $\ell \in \{1,\ldots,L-1\}$, we have $\rho^2_{(\boldsymbol{X}^M_{(\ell)}\to Y)} = \rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} = 0$, and hence $r_{V|\boldsymbol{U}^M_{(\ell)}}(\boldsymbol{u}^M_{(\ell)}) = r_{V|\boldsymbol{U}^Q_{(\ell)}}(\boldsymbol{u}^Q_{(\ell)}) = r_{V|\boldsymbol{U}^Q_{(\ell)},\boldsymbol{U}^M_{(\ell+1)}}(\boldsymbol{u}^Q_{(\ell)}, \boldsymbol{u}^M_{(\ell+1)}) = 1/2$ for any $\boldsymbol{u}^Q_{(\ell)}$ and $\boldsymbol{u}^M_{(\ell+1)}$.

(ii) If there is a ℓ^* with $1 \le \ell^* \le L-1$ such that

$$\begin{aligned}
\rho^2_{(\boldsymbol{X}\to Y)} &= \rho^2_{(\boldsymbol{X}^M_{(1)}\to Y)} && \text{for} \quad \ell^* = 1,\\
&= \rho^2_{(\boldsymbol{X}^M_{(1)}\to Y)} + \sum_{\ell'=1}^{\ell^*-1} \rho^2_{(\boldsymbol{X}^M_{(\ell'+1)}\to Y|\boldsymbol{X}^Q_{(\ell')})} && \text{for} \quad 2 \le \ell^* \le L-1,
\end{aligned}$$

then for every $\ell \in \{\ell^*,\ldots,L-1\}$, we have $\rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} = 0$ and hence $r_{V|\boldsymbol{U}^Q_{(\ell)},\boldsymbol{U}^M_{(\ell+1)}}(\boldsymbol{u}^Q_{(\ell)}, \boldsymbol{u}^M_{(\ell+1)}) = r_{V|\boldsymbol{U}^Q_{(\ell)}}(\boldsymbol{u}^Q_{(\ell)})$ for any $\boldsymbol{u}^Q_{(\ell)}$ and $\boldsymbol{u}^M_{(\ell+1)}$.

Proof. See Appendix A.3. □

According to Proposition 6(i), $\rho^2_{(\boldsymbol{X}\to Y)} = 0$ implies that none of d independent variables has contribution to the construction of the mean checkerboard copula score in Y with respect to any conditional distribution of Y given independent variables in $\boldsymbol{X}$. Proposition 6(ii) indicates that if the condition holds, then the independent variables in $\boldsymbol{X}^M_{(\ell+1)}$ for every $\ell \in \{\ell^*,\ldots,L-1\}$ have no additional contribution to the computation of the mean checkerboard copula score of Y given $\boldsymbol{X}^Q_{(\ell^*)}$.

As explained in Section 2, the upper bound of the overall CCRAM $\rho^2_{(\boldsymbol{X}\to Y)}$ is $12\sigma^2_Y$ which depends on the marginal distribution of Y equipped with its checkerboard copula score, $\{s_j\}_{j=1}^J$. Thus, it is natural to consider the marginal/conditional Scaled CCRAM(SCCRAM)s and the corresponding sequential decomposition of the overall SCCRAM :

$$\rho^{2\,*}_{(\boldsymbol{X}\to Y)} = \rho^{2\,*}_{(\boldsymbol{X}^M_{(1)}\to Y)} + \sum_{\ell=1}^{L-1} \rho^{2\,*}_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})}, \tag{15}$$

where $\rho^{2\,*}_{(\boldsymbol{X}\to Y)} = \rho^2_{(\boldsymbol{X}\to Y)}\big/12\sigma^2_Y$, $\rho^{2\,*}_{(\boldsymbol{X}^M_{(1)}\to Y)} = \rho^2_{(\boldsymbol{X}^M_{(1)}\to Y)}\big/12\sigma^2_Y$ and $\rho^{2\,*}_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} = \rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})}\big/12\sigma^2_Y$ are the overall, marginal, and conditional SCCRAMs. It is also worth pointing out that the properties shown in Propositions 3–6 and Theorem 2(b) still hold for the overall/marginal/conditional SCCRAMs.

4. Statistical inference

To estimate the overall CCRAM/SCCRAM reviewed in Section 2, the marginal/conditional CCRAMs/SCCRAMs, and the decomposition of the overall CCRAM/SCCRAM proposed in Section 3, we use the plug-in method based on the relative frequencies of a multidimensional contingency table. Let $\{n_{\boldsymbol{i},j}\}$, $\boldsymbol{i} = (i_1, \ldots, i_k, \ldots, i_d)$, $i_k \in \{1, \ldots, I_k\}$, $k \in \{1, \ldots, d\}$, $j \in \{1, \ldots, J\}$, denote the cell counts in an observed contingency table obtained by classifying n cases according to categories of a $(d+1)$-dimensional ordinal vector $(\mathbf{X}^T, Y)$ where $\boldsymbol{X} = (X_1, \ldots, X_k, \ldots, X_d)$. We then have the relative frequency $\left\{\hat{p}_{\boldsymbol{i},j} = n_{\boldsymbol{i},j}/n\right\}$ as an estimator of the joint probability mass function of $(\mathbf{X}^T, Y)$, $\{p_{\boldsymbol{i},j}\}$ where $n = \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{I}} \sum_{j=1}^{J} n_{\boldsymbol{i},j}$, $\mathbf{1} = (1, \ldots, 1)$ is of length d, and $\boldsymbol{I} = (I_1, \ldots, I_d)$. We also obtain the estimators for the p.m.f. of d-dimensional marginal table for $\mathbf{X}$, the p.m.f.s of 1-dimensional marginal tables for X_k and Y, and the conditional p.m.f. for Y given $\mathbf{X}$, respectively:

$$\hat{p}_{\boldsymbol{i}+} = \frac{n_{\boldsymbol{i}+}}{n}, \quad \hat{p}_{+i_k+} = \frac{n_{+i_k+}}{n}, \quad \hat{p}_{+j} = \frac{n_{+j}}{n}, \quad \hat{p}_{j|\boldsymbol{i}} = \frac{\hat{p}_{\boldsymbol{i},j}}{\hat{p}_{\boldsymbol{i}+}},$$

where $n_{\boldsymbol{i}+}$, n_{+i_k+} and n_{+j} are obtained by summing $n_{\boldsymbol{i},j}$ over all levels in Y, levels of all other factors except for X_k and Y, respectively. Furthermore, the ranges of marginal c.d.f.s of X_k and Y are estimated by $\hat{D}_{X_k} = \{\hat{u}_0^k, \ldots, \hat{u}_{i_k}^k, \ldots, \hat{u}_{I_k}^k\}$ and $\hat{D}_Y = \{\hat{v}_0, \ldots, \hat{v}_j, \ldots, \hat{v}_J\}$ where $\hat{u}_{i_k}^k = \hat{F}_{X_k}(x_{i_k}^k) = \sum_{\ell_k=1}^{i_k} \hat{p}_{+\ell_k+}$, $\hat{v}_j = \hat{F}_Y(y_j) = \sum_{\ell=1}^{j} \hat{p}_{+\ell}$, $\hat{u}_0^k = \hat{v}_0 = 0$, and $\hat{u}_{I_k}^k = \hat{v}_J = 1$.

By utilizing the estimators above, we can obtain the plug-in estimators for the overall, marginal and conditional CCRAMs in (4), (11) and (12), the sequential decomposition of the overall CCRAM in (14), and the corresponding SCCRAMs in (15):

$$\hat{\rho}^2_{(\boldsymbol{X}\to Y)} = 12 \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{I}} \left(\sum_{j=1}^{J} \hat{s}_j \hat{p}_{j|\boldsymbol{i}} - 1/2 \right)^2 \hat{p}_{\boldsymbol{i}+}, \quad \hat{\rho}^{2\,*}_{(\boldsymbol{X}\to Y)} = \frac{\hat{\rho}^2_{(\boldsymbol{X}\to Y)}}{12\hat{\sigma}_Y^2},$$

$$\hat{\rho}^2_{(\boldsymbol{X}^Q_{(\ell)}\to Y)} = 12 \sum_{\boldsymbol{i}^Q_{(\ell)}=\mathbf{1}}^{\boldsymbol{I}^Q_{(\ell)}} \left(\sum_{j=1}^{J} \hat{s}_j \hat{p}_{j|\boldsymbol{i}^Q_{(\ell)}} - 1/2 \right)^2 \hat{p}_{\boldsymbol{i}^Q_{(\ell)}+}, \quad \hat{\rho}^{2\,*}_{(\boldsymbol{X}^Q_{(\ell)}\to Y)} = \frac{\hat{\rho}^2_{(\boldsymbol{X}^Q_{(\ell)}\to Y)}}{12\hat{\sigma}_Y^2},$$

$$\hat{\rho}^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} = 12 \sum_{\boldsymbol{i}^Q_{(\ell)}=\mathbf{1}}^{\boldsymbol{I}^Q_{(\ell)}} \sum_{\boldsymbol{i}^M_{(\ell+1)}=\mathbf{1}}^{\boldsymbol{I}^M_{(\ell+1)}} \left(\sum_{j=1}^{J} \hat{s}_j \hat{p}_{j|\boldsymbol{i}^Q_{(\ell)},\boldsymbol{i}^M_{(\ell+1)}} - \sum_{j=1}^{J} \hat{s}_j \hat{p}_{j|\boldsymbol{i}^Q_{(\ell)}} \right)^2 \hat{p}_{\boldsymbol{i}^Q_{(\ell)},\boldsymbol{i}^M_{(\ell+1)}+}, \quad \hat{\rho}^{2\,*}_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} = \frac{\hat{\rho}^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})}}{12\hat{\sigma}_Y^2},$$

$$\hat{\rho}^2_{(\boldsymbol{X}\to Y)} = \hat{\rho}^2_{(\boldsymbol{X}^M_{(1)}\to Y)} + \sum_{\ell=1}^{L-1} \hat{\rho}^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})}, \quad \hat{\rho}^{2\,*}_{(\boldsymbol{X}\to Y)} = \hat{\rho}^{2\,*}_{(\boldsymbol{X}^M_{(1)}\to Y)} + \sum_{\ell=1}^{L-1} \hat{\rho}^{2\,*}_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})},$$

where $\hat{s}_j = (\hat{v}_{j-1} + \hat{v}_j)/2$ and $\hat{\sigma}_Y^2 = \sum_{j=1}^{J} \hat{v}_{j-1}\hat{v}_j/4\hat{p}_{+j}$. Note that the sequential decomposition of the overall CCRAM and its SCCRAM hold exactly for the plug-in estimators of the overall, marginal and conditional CCRAMs and their SCCRAMs.

Note that to properly evaluate the magnitudes of the associations measured by the overall/marginal/conditional CCRAMs, we will present the results of the overall SCCRAM and its sequential decomposition (based on marginal/conditional SCCRAMs) in the data analysis section (see Section 5).

For the statistical significance of different types of regression associations measured by the overall/marginal/ conditional SCCRAMs, we propose using the permutation tests under the null hypotheses [7,8,24] to obtain p-values associated with the estimated SCCRAMs. When we perform multiple hypotheses testing involving marginal/conditional SCCRAMs in the decomposition of the overall SCCRAM, the Benjamini–Yekutieli FDR-controlling method [6] is applied to adjust the associated p-values for controlling the false discovery rate.

In practice, we may expect some degree of association even if it is small. In such a case, the interval estimation can be a useful tool if one wishes to learn more from estimating the magnitude of a SCCRAM of interest. Thus, we propose computing the bias-corrected and accelerated (BCa) bootstrap interval obtained from the nonparametric bootstrap method (i.e., B bootstrap samples of size n from the observed contingency table with cell counts $\{n_{\boldsymbol{i},j}\}$) [11,17]. Moreover, the bootstrap approach can take into account the finite sample variation in the estimation of (overall/marginal/conditional) SCCRAMs and also preserve the theoretical range of the SCCRAMs to be between 0 and 1. It is worthy to point out that the bootstrap confidence intervals that will be presented in the case studies of Section 5 are individual (not simultaneous) confidence intervals.

Remark 1. The permutation test described above will be performed under the null hypothesis that there is no regression association between the dependent ordinal variable and a set of categorical independent variables of interest. However, the bootstrap samples will be generated from the observed data and we will obtain the corresponding confidence interval around an estimated SCCRAM. Furthermore, the theoretical lower bound of (overall/marginal/conditional) SCCRAMs is 0. Therefore, the bootstrap confidence interval may not be a suitable tool for testing the hypothesis that a SCCRAM of interest is 0.

Table 1
Responses from study participants cross-classified with respect to fatigue (F), weakness (W), time point (TIME) and treatment (TRT) in a four-dimensional contingency table.

TIME	F	TRT	Control				Intervention			
		W	Absent	Mild	Moderate	Severe	Absent	Mild	Moderate	Severe
Start	absent		53	10	6	6	83	11	6	5
Start	mild		10	35	7	0	15	39	13	1
Start	moderate		8	18	46	12	16	28	41	4
Start	severe		3	8	17	22	4	6	19	24
1 Month	absent		45	5	5	4	69	9	6	3
1 Month	mild		10	22	6	0	11	18	5	0
1 Month	moderate		10	14	37	4	10	14	22	3
1 Month	severe		4	1	4	17	6	4	13	12
3 Month	absent		47	9	2	5	87	5	7	2
3 Month	mild		13	24	4	1	11	12	6	2
3 Month	moderate		15	12	22	4	12	12	4	1
3 Month	severe		2	0	5	17	3	3	3	12

Remark 2. For hypothesis testing of no (overall/marginal/conditional) regression association, one may consider deriving the asymptotic distribution of the estimator for the corresponding SCCRAM under the null hypothesis. Although the asymptotic tests are more computationally efficient than the permutation tests, however, there are several advantages for the latter [7,8]. First, like the estimation of the SCCRAMs, the permutation test is entirely data-dependent in that all the required information for analysis is based on the observed data only. Second, in the same spirit of the proposed SCCRAMs, the permutation test is distribution-free because it does not require a specific distribution for the population from which the sample was drawn or any parametric assumptions on the association structure. Third, unlike the permutation test, the asymptotic test could be problematic when the sample size is not sufficiently large compared to the size of contingency table (e.g., for sparse contingency tables) [3,29]. Furthermore, for small data sets, the sampling distribution of the test statistic of interest is likely to be discrete and an approximating theoretical distribution tends to give poor fits to the underlying discrete sampling distribution [8, p. 60].

5. Case studies

In this section we will illustrate how the proposed model-free methods enable identifying and quantifying different aspects of association structures for multivariate categorical data with ordinal response variable(s) in an exploratory and meaningful way. To this end, we consider two real data sets analyzed in the literature, the first one from a randomized, controlled trial of the progression of fatigue and weakness in HIV-positive individuals [13] and the second one from the longitudinal epidemiologic study of post-traumatic stress after the Three Mile Island accident [19]. For both data sets, model-based approaches were proposed by researchers and asymptotic tests were used for the goodness of fit of the models and statistical significance of the associations between the variables in the fitted models. We will first apply the proposed methods to these two real datasets and see if the proposed methods can provide consistent information on association structures as those by the model-based approaches in the literatures.

5.1. Fatigue and weakness in people living with HIV

Weakness and fatigue are two known symptoms for People Living with HIV (PLHIV) and they are closely related to each other when self-rated by PLHIV [13,16,20]. Table 1 is a four-dimensional contingency table obtained from a recent longitudinal randomized controlled clinical trial of HIV-positive individuals. The study participants were cross-classified by four variables: *fatigue* (F), *weakness* (W), *treatment* (TRT) and *data collection time* ($TIME$). The variables *fatigue* and *weakness*, self scored by the study participants, are ordinal with four levels ("absent < "mild" < "moderate" < "severe"), the *time* is also ordinal with three levels ("start of the study", "one month", "three months"), and the *treatment* is a nominal variable with two levels ("control" group, "intervention" group). Dobra et al. [13] illustrated the application of hierarchical, graphical log-linear models to this data set to investigate the causal pathway through which the treatment affects the degree of fatigue and weakness over time and to determine mediating variables in the causal process.

In order to properly carry out a model-based causal pathway analysis, it is imperative to explore the data set of interest to detect its multivariate association structure and distill meaningful insights for data-driven statistical modeling. For example, in Table 1, one main interest might be to identify and quantify the pairwise association between W and F, with or without conditioning on $TIME$ and TRT. Another interest could be to discover empirical evidence for the absence/presence of regression association, conditional or not, among the variables by asking questions such as "is there a regression effect of $TIME$ and TRT on F (W) without or with conditioning on W (F)?".

Table 2 shows the results of using the overall SCCRAM and its decompositions to measure the global/marginal/conditional regression associations among the four variables in Table 1. Two types of overall regression dependence structure are considered. The first type uses W as a dependent variable and ($TIME$, TRT, F) as independent variables,

Table 2
Estimates of the overall, marginal and conditional SCCRAMs for two types of regression structure $(TIME, TRT, F) \rightarrow W$ and $(TIME, TRT, W) \rightarrow F$ in the HIV data of Table 1, along with the contribution percentage (%) of the marginal and conditional SCCRAMs to the overall SCCRAM, the individual bootstrap confidence intervals and the (adjusted) permutation p-values.

Overall SCCRAM	Decomposition	Estimate	Contribution (%)	95% BCa bootstrap C.I.	(Adjusted) Permutation *p*-value
$\rho^{2\ *}_{(TIME,TRT,F)\rightarrow W}$ $(\hat{\sigma}^2_W = 0.903)$		0.388		(0.321, 0.426)	0.0000
	$\rho^{2\ *}_{F\rightarrow W}$	0.366	94.3(=100*0.366/0.388)	(0.312, 0.412)	0.0000
	$\rho^{2\ *}_{(TIME,TRT)\rightarrow W\mid F}$	0.022	5.7(=100*0.022/0.388)	(0.014, 0.025)	0.0013
	$\rho^{2\ *}_{(TIME,TRT)\rightarrow W}$	0.038	10(=100*0.038/0.388)	(0.022, 0.054)	0.0000
	$\rho^{2\ *}_{F\rightarrow W\mid(TIME,TRT)}$	0.350	90(=100*0.350/0.388)	(0.284, 0.389)	0.0000
$\rho^{2\ *}_{(TIME,TRT,W)\rightarrow F}$ $(\hat{\sigma}^2_F = 0.917)$		0.377		(0.320, 0.402)	0.0000
	$\rho^{2\ *}_{W\rightarrow F}$	0.365	96.8 (=100*0.365/0.377)	(0.313, 0.414)	0.0000
	$\rho^{2\ *}_{(TIME,TRT)\rightarrow F\mid W}$	0.012	3.2 (=100*0.012/0.377)	(0.008, 0.012)	0.3846
	$\rho^{2\ *}_{(TIME,TRT)\rightarrow F}$	0.026	6.9 (=100*0.026/0.377)	(0.013, 0.038)	0.0000
	$\rho^{2\ *}_{W\rightarrow F\mid(TIME,TRT)}$	0.351	93.1(=100*0.351/0.377)	(0.294, 0.392)	0.0000

denoted as $(TIME, TRT, F) \rightarrow W$, and the second type is $(TIME, TRT, W) \rightarrow F$ where F is considered as a dependent variable and $(TIME, TRT, W)$ as independent variables. The overall association for each regression structure is 0.388 for the first regression structure and 0.377 for the second one, implying that the average proportions of variance for W and F explained by the checkerboard copula regression using $(TIME, TRT, F)$ and $(TIME, TRT, W)$ are 38.8% and 37.7%, respectively. The permutation p-values very close to zero for $H_0 : \rho^{2\ *}_{(TIME,TRT,F)\rightarrow W} = 0$ and $H_0 : \rho^{2\ *}_{(TIME,TRT,W)\rightarrow F} = 0$ indicate that the set of the independent variables considered in each regression structure has statistically significant contribution to the variance for the dependent variable.

Moreover, several observations are revealed in Table 2 through the proposed sequential decomposition. First, F and W are directly associated with each other, with or without conditioning on $TIME$ and TRT because the magnitudes of association between F and W are all between 0.35 and 0.37, $(\rho^{2\ *}_{F\rightarrow W}, \rho^{2\ *}_{F\rightarrow W\mid(TIME,TRT)})$=(0.366, 0.350) and $(\rho^{2\ *}_{W\rightarrow F}, \rho^{2\ *}_{W\rightarrow F\mid(TIME,TRT)})$=(0.365, 0.351), and they are all statistically significant. It is worth noting that the percentages of the (marginal/ conditional) contribution of F and W to the global association in each regression structure range from 90% to 96.8% (see the column "Contribution (%)" in Table 2). Second, $(TIME, TRT)$ has a significant marginal effect on W and F, respectively, according to the permutation p-values, but their magnitudes, $\rho^{2\ *}_{(TIME,TRT)\rightarrow W}$=0.038 and $\rho^{2\ *}_{(TIME,TRT)\rightarrow F}$=0.026, are small relative to the overall regression associations. Third, while the regression effect of $TIME$ and TRT on W conditional on F, $\rho^{2\ *}_{(TIME,TRT)\rightarrow W\mid F}$=0.022, is still significant, it appears that, conditionally on W, there is no significant regression association between $(TIME, TRT)$ and F $(\rho^{2\ *}_{(TIME,TRT)\rightarrow F\mid W}$=0.012). This last finding tells us that while HIV patients' treatment (TRT) and $TIME$ would provide some insights about their weakness level (W), even if we knew of their fatigue level (F), the other way around does not hold; in other words, given the knowledge of a HIV patient's weakness level, knowing which treatment they are receiving and how long they have been treated in fact gives us very little information about their fatigue level.

All the results shown above, especially the first and third findings from the sequential decomposition, echo well the results found by Dobra et al. [13] using the hierarchical graphical log-linear model with minimal sufficient statistics $[W, F][W, TIME][W, TRT]$ representing the model of conditional independence of $TIME$, TRT and F given W. Note that the conditional independence log linear model was selected by the asymptotic log-likelihood ratio tests and their associated p-values.

5.2. Three Mile Island stress-level study

The Three Mile Island (TMI) data, shown in Table 3, is the repeatedly measured categorical data obtained from a longitudinal epidemiological study of the mental health effects associated with the Three Mile Island accident occurred in 1979 [19]. As this study focused on the mental health of mothers of young children living within ten miles of the plant, the four waves of interviews were conducted to measure the post accident stress level (summarized from a 90-item self-report checklist) of the participants in Winter 1979, Spring 1980, Fall 1981, and Fall 1982; as a risk factor of the psychological stress, the distance of residence from the plant was also recorded. The responses from 267 mothers were cross-classified with respect to the stress level at each wave, denoted as W_1, W_2, W_3, W_4, each with three ordinal levels (Low, Medium, High) and the residence distance, denoted as D, with two levels (<5 miles vs. > 5 miles), in a five-dimensional contingency table as shown in Table 3.

Several researchers have analyzed this TMI data set using various types of model-based method, including additive continuation ratio logit models and semiparametric mixed logistic regression models [19], mixed model approach via conditional likelihood approach [9], linear dynamic conditional multinomial probability and multinomial dynamic logit models [51, Sec. 3.4], and discretized parametric copula models with the univariate margins estimated from empirical counterparts [31, Sec. 1.5.4].

Table 3
Post accident stress levels of mothers of young children living within ten miles of the Three Mile Island nuclear power plant measured at four time points after the nuclear accident.

D			< 5 miles			> 5 miles			D			< 5 miles			> 5 miles		
			W_4									W_4					
W_1	W_2	W_3	L	M	H	L	M	H	W_1	W_2	W_3	L	M	H	L	M	H
L	L	L	2	0	0	1	2	0	M	M	H	0	2	3	0	5	1
L	L	M	2	3	0	2	0	0	M	H	L	0	0	0	0	0	0
L	L	H	0	0	0	0	0	0	M	H	M	0	2	0	0	1	1
L	M	L	0	1	0	1	0	0	M	H	H	0	1	1	0	3	1
L	M	M	2	4	0	0	3	0	H	L	L	0	0	0	0	0	1
L	M	H	0	0	0	0	0	0	H	L	M	0	0	0	0	0	0
L	H	L	0	0	0	0	0	0	H	L	H	0	0	0	0	0	0
L	H	M	0	0	0	0	0	0	H	M	L	0	0	0	0	0	0
L	H	H	0	0	0	0	0	0	H	M	M	0	4	3	1	13	0
M	L	L	5	1	0	4	4	0	H	M	H	0	1	4	0	0	0
M	L	M	1	4	0	5	15	1	H	H	L	0	0	0	0	0	0
M	L	H	0	0	0	0	0	0	H	H	M	1	2	0	0	7	2
M	M	L	3	2	0	2	2	0	H	H	H	0	5	12	0	2	7
M	M	M	2	38	4	6	53	6									

Table 4
Estimates of overall/marginal/conditional SCCRAMs for two regression structures, $(W_1, W_2, W_3) \to W_4$ (Case I) and $(D, W_1, W_2, W_3) \to W_4$ (Case II) in TMI data of Table 3 along with the corresponding individual bootstrap confidence intervals and the (adjusted) permutation p-values. Note that Case II considers Distance (D) as an independent variable or a conditioning factor while Case I does not.

Case	Overall SCCRAM	Marginal/Conditional SCCRAM	Estimate	95% BCa bootstrap C.I.	(Adjusted) Permutation p-value
I : without Distance (D)					
	$\rho^{2\,*}_{(W_1,W_2,W_3)\to W_4}$		0.4055	(0.2812, 0.4678)	0.0000
		$\rho^{2\,*}_{W_1\to W_2}$	0.3173	(0.2229, 0.3998)	0.0000
		$\rho^{2\,*}_{W_2\to W_3}$	0.3635	(0.2641, 0.4491)	0.0000
		$\rho^{2\,*}_{W_3\to W_4}$	0.3094	(0.2216, 0.3975)	0.0000
		$\rho^{2\,*}_{W_1\to W_3}$	0.1973	(0.1256, 0.2632)	0.0000
		$\rho^{2\,*}_{W_2\to W_4}$	0.2130	(0.1367, 0.2933)	0.0000
		$\rho^{2\,*}_{W_1\to W_4}$	0.1846	(0.1121, 0.2578)	0.0000
		$\rho^{2\,*}_{W_2\to W_3\mid W_1}$	0.1874	(0.1178, 0.2539)	0.0000
		$\rho^{2\,*}_{W_3\to W_4\mid(W_1,W_2)}$	0.1351	(0.0612, 0.1854)	0.0000
		$\rho^{2\,*}_{W_1\to W_3\mid W_2}$	0.0213	(0.0061, 0.0386)	0.4269
		$\rho^{2\,*}_{W_2\to W_4\mid(W_1,W_3)}$	0.0201	(0.0042, 0.0265)	1.0000
		$\rho^{2\,*}_{W_1\to W_4\mid(W_2,W_3)}$	0.0635	(0.0377, 0.0802)	0.0408
II : with Distance (D)					
	$\rho^{2\,*}_{(D,W_1,W_2,W_3)\to W_4}$		0.4624	(0.3910, 0.4950)	0.0000
		$\rho^{2\,*}_{W_1\to W_2\mid D}$	0.3142	(0.2179, 0.4003)	0.0000
		$\rho^{2\,*}_{W_2\to W_3\mid D}$	0.3639	(0.2645, 0.4598)	0.0000
		$\rho^{2\,*}_{W_3\to W_4\mid D}$	0.3174	(0.2104, 0.3986)	0.0000
		$\rho^{2\,*}_{W_1\to W_3\mid D}$	0.2325	(0.1533, 0.3027)	0.0000
		$\rho^{2\,*}_{W_2\to W_4\mid D}$	0.2245	(0.1511, 0.2954)	0.0000
		$\rho^{2\,*}_{W_1\to W_4\mid D}$	0.1944	(0.1329, 0.2616)	0.0000
		$\rho^{2\,*}_{W_2\to W_3\mid(D,W_1)}$	0.1826	(0.1066, 0.2373)	0.0000
		$\rho^{2\,*}_{W_3\to W_4\mid(D,W_1,W_2)}$	0.1513	(0.0710, 0.1949)	0.0000
		$\rho^{2\,*}_{W_1\to W_3\mid(D,W_2)}$	0.0512	(0.0108, 0.0806)	0.0536
		$\rho^{2\,*}_{W_2\to W_4\mid(D,W_1,W_3)}$	0.0631	(0.0402, 0.0745)	0.2011
		$\rho^{2\,*}_{W_1\to W_4\mid(D,W_2,W_3)}$	0.0805	(0.0462, 0.0960)	0.0461

A main interest of the TMI data in the model-building process is to study how stress level of the subjects changed over time with purposes of uncovering time-dependent association among the four waves and providing insights on the follow-up modeling steps. In particular, it would be useful to understand how the stress levels at previous waves contributed to the stress level at a later wave of interest, while conditional on previous and intervening waves and/or time-independent predictor, the residence distance (D).

Table 4 summarizes various aspects of regression association for TMI data in Table 3 measured by the overall/marginal/conditional SCCRAMs. In order to take into account time ordering in computing the proposed SCCRAMs, the

(a) Case I: no conditioning on D ($\rho^{2\,*}_{(W_1,W_2,W_3)\to W_4} = 0.4055$) (b) Case II: conditioning on D ($\rho^{2\,*}_{(D,W_1,W_2,W_3)\to W_4} = 0.4624$)

Fig. 1. Time-dependent association structure for TMI data, without (left) and with (right) conditioning on Distance. Note that the numbers outside(inside) the parenthesis at each edge denote the values of the regression association between two connected waves without (with) conditioning on previous and intervening waves, and the red bold numbers indicate significant regression associations according to the permutation tests (p-value<0.01). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

stress levels measured at a later wave and an earlier wave are considered as a dependent variable and an independent variable. Fig. 1 visualizes the time-dependent association structure shown in Table 4.

From Table 4 and Fig. 1(a), we first examine the dependence structure for TMI data without considering the distance (D), denoted as Case I. The significant overall regression dependence between (W_1, W_2, W_3) and W_4 with $\rho^{2\,*}_{(W_1,W_2,W_3)\to W_4}$= 0.4055 indicates that the average proportion of variance for the checkerboard copula score of W_4 explained by the checkerboard copula regression using (W_1, W_2, W_3) is 40.55%. Next, the marginal associations between two waves, consecutive or not (denoted as $\rho^{2\,*}_{W_i\to W_j}$ where $i < j$ and $i, j \in \{1, 2, 3, 4\}$), are all statistically significant and the strength of regression associations decreases as the time lag $(j - i)$ between waves increases. In particular, the magnitude of the association between two consecutive waves $(j - i = 1)$ is at least 1.45 times larger than those between waves whose time lag is 2 and 3 (e.g., $\rho^{2\,*}_{W_3\to W_4}/\rho^{2\,*}_{W_2\to W_4}$ = 0.3094/0.2130 = 1.45).

However, when the stress levels at all previous and intervening waves are taken into account in computing the conditional association between the stress levels at two waves of interest (denoted as $\rho^{2\,*}_{W_i\to W_j|\text{ all remaining } W_k}$ in Table 4 where $k \in \{1, \ldots, 4\}$, $k \neq i$ and $k < j$), the behavior of conditional associations is quite different from that of marginal associations (see the numbers inside parentheses in Fig. 1(a) as well). Specifically, the magnitudes of the conditional regression associations are reduced by half or more, compared to the corresponding marginal associations; for instance, $\rho^{2\,*}_{W_2\to W_3|W_1}$=0.1874 is half of $\rho^{2\,*}_{W_2\to W_3}$=0.3635, while $\rho^{2\,*}_{W_1\to W_3|W_2}$=0.0213 is merely tenth of $\rho^{2\,*}_{W_1\to W_3}$=0.1973. Moreover, the magnitudes of the conditional association between the non-consecutive waves are not significantly different from zero (corresponding p-values$>$ 0.01) and is less than half of those between two consecutive waves, which are statistically significant.

When the distance (D) is considered, denoted as Case II in Table 4 and Fig. 1(b), we observe similar patterns of associations except that (i) the magnitude of the overall association increases by about 6% and (ii) the other association measures tend to be similar to or slightly larger than those without conditioning on the distance.

According to the results shown above, we make the following four exploratory findings. First, the marginal pairwise associations between two waves without or with conditioning on the distance decrease as the time lag between waves increases. This was also found by the model-based approaches in [51, Sec. 3.4] and [31, Sec. 1.5.4]. Second, it appears that the distance does not have any essential effect on the pattern of the association between two waves of interest whether or two waves are consecutive or all previous and intervening waves are accounted for. Third, the magnitude of the regression dependence between the stress levels at two waves of interest does depend on the stress levels at the previous and intervening waves. For example, $\rho^{2\,*}_{W_3\to W_4}$ and $\rho^{2\,*}_{W_3\to W_4|D}$ are at least twice larger than $\rho^{2\,*}_{W_3\to W_4|(W_1,W_2)}$ and $\rho^{2\,*}_{W_3\to W_4|(D,W_1,W_2)}$, respectively, which means that the stress levels at W_1 and W_2 have psychological residual effect on the stress level at W_4. Note that the second and third observations above are consistent with those found in [19]. Fourth, the dependence between the stress levels at two waves adjusted for those at the previous and intervening waves appears to have a first-order Markov structure, as the conditional associations between consecutive waves are significant while those between non-consecutive waves are not.

6. Discussion

In this paper we propose a data-analytic approach to exploring and uncovering various characteristics of association structure for the multivariate categorical data with an ordinal response variable. The proposed approach is based on decomposing the global regression association between an ordinal response variable and all available categorical independent variables (measured by the overall CCRAM and SCCRAM) into a marginal regression association and several conditional regression associations quantified by the proposed marginal and conditional CCRAMs and SCCRAMs. The proposed decomposition is sequential in nature in that the global regression association can be partitioned in various ways depending on the order of the independent variables. Thus, it can allow us to quantify marginal and conditional contributions of any subset of categorical independent variables to the overall regression association, without or with conditional on other independent variables, and it can provide insightful and plentiful information on complex regression dependence structures.

As a reviewer pointed out, it would be desirable to investigate the finite-sample performances of the estimators of the overall/marginal/conditional SCCRAMs and the associated permutation tests. To this end, we have conducted simulation studies to examine (i) the unbiasedness, variability, sampling distribution and empirical convergence rate of the estimators for the SCCRAMs and (ii) the Type I error rates/sizes and powers of the permutation tests in a three-way contingency table with an ordinal dependent variable and two ordinal independent variables.

Due to the limited space, we briefly summarize the findings from the simulation studies given in the supplementary material. First, the bias and variability of the estimators for the overall/marginal/conditional SCCRAMs tend to decrease as the sample size n increases, regardless of the table size (number of categories in the variable) and the magnitude of the SCCRAM. Note that the rate at which bias is attenuated appears to differ, depending on whether or not the true value of the SCCRAM is bounded away from 0. Second, when the true values for the SCCRAMs are away from 0, the sampling distributions of the corresponding estimators appear to be symmetric and the empirical convergence rates of the estimators are approximately of order $n^{-0.5}$, irrespective of the table size. Furthermore, we observe that the densities of the normalized (centered and rescaled) estimators of the SCCRAMs appear to be close to a normal density with zero mean. Thus, it is anticipated that the asymptotic distributions of the estimators of the proposed SCCRAMs would be normally distributed provided that there is a sufficiently large sample size and the true value is bounded away from 0. However, when the true value is exactly 0 or close to 0 (0.05 and below) and n is small, the sampling distributions of the estimators appear to be skewed to the right. Third, the permutation tests for overall/marginal/conditional SCCRAMs have a good control of Type I error rates in the sense that the sizes of the tests tend to be fairly close to the nominal significance levels. The powers of the permutation tests for the SCCRAMs increase and approach to 1 as either n increases or the magnitude of marginal/conditional SCCRAMs departs from 0 and increases. For detailed information on the simulation (simulation scheme and simulation results), please see Sections S1 and S2 in the supplementary material.

In the real data examples we utilized resampling methods for the interval estimation and the hypothesis testing associated with the estimators of various forms of marginal and conditional regression association measures in the decomposition of the overall regression association measure. In practice, these methods are suitable when the sample size is not large relative to the size of the contingency table (i.e., number of variables and number of categories of each variable). A potential future extension of this research would be to investigate the asymptotic distributions of the estimators of the marginal and conditional regression association measures for interval estimation and hypothesis testing.

CRediT authorship contribution statement

Zheng Wei: Methodology, Software, Formal analysis, Visualization, Writing – original draft, Writing editing. **Li Wang:** Methodology, Software, Validation, Formal analysis, Visualization, Writing – original draft. **Shu-Min Liao:** Conceptualization, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Daeyoung Kim:** Conceptualization, Methodology, Formal analysis, Validation, Visualization, Supervision, Writing – original draft, Writing – review & editing.

Acknowledgments

The authors would like to thank the Editor, Associate Editor and two anonymous referees for their insightful and constructive comments and suggestions, which helped us greatly in revising this work.

Appendix A. Proofs

The proof of Theorem 2, Propositions 3 and 4. It is sufficient to show that $\rho^2_{(\boldsymbol{X}\to Y)} = \rho^2_{(\boldsymbol{X}^Q_{(L-1)}\to Y)} + \rho^2_{(\boldsymbol{X}^M_{(L)}\to Y|\boldsymbol{X}^Q_{(L-1)})}$ for $\boldsymbol{X}^Q_{(L-1)}$ and $\boldsymbol{X}^M_{(L)}$ in a partition M of $\boldsymbol{X}$, where $\boldsymbol{X}^Q_{(L-1)}$ and $\boldsymbol{X}^M_{(L)}$ are defined at the beginning of Section 3.2. Then, $\rho^2_{(\boldsymbol{X}^Q_{(L-1)}\to Y)}, \ldots, \rho^2_{(\boldsymbol{X}^Q_{(2)}\to Y)}$ can be recursively decomposed in the same way as we decompose $\rho^2_{(\boldsymbol{X}\to Y)}$. In fact, according to definitions of $\rho^2_{(\boldsymbol{X}\to Y)}$ in (4), $\rho^2_{(\boldsymbol{X}^Q_{(L-1)}\to Y)}$ in (11), and $\rho^2_{(\boldsymbol{X}^M_{(L)}\to Y|\boldsymbol{X}^Q_{(L-1)})}$ in (12), we have

$$\rho^2_{(\boldsymbol{X}\to Y)} = 12\mathrm{Var}\big[r_{V|\boldsymbol{U}}(\boldsymbol{U})\big], \quad \rho^2_{(\boldsymbol{X}^Q_{(L-1)}\to Y)} = 12\mathrm{Var}\Big[r_{V|\boldsymbol{U}^Q_{(L-1)}}\Big(\boldsymbol{U}^Q_{(L-1)}\Big)\Big],$$

$$\rho^2_{(\boldsymbol{X}^M_{(L)}\to Y|\boldsymbol{X}^Q_{(L-1)})} = 12E\left[\left\{r_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}\Big(\boldsymbol{U}^Q_{(L-1)}, \boldsymbol{U}^M_{(L)}\Big) - r_{V|\boldsymbol{U}^Q_{(L-1)}}\Big(\boldsymbol{U}^Q_{(L-1)}\Big)\right\}^2\right],$$

where

$$r_{V|\boldsymbol{U}^Q_{(L-1)}}\Big(\boldsymbol{U}^Q_{(L-1)}\Big) = \mathrm{E}_{c^+_{V|\boldsymbol{U}^Q_{(L-1)}}}\Big(V|\boldsymbol{U}^Q_{(L-1)}\Big), \quad r_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}\Big(\boldsymbol{U}^Q_{(L-1)}, \boldsymbol{U}^M_{(L)}\Big) = \mathrm{E}_{c^+_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}}\Big(V|\boldsymbol{U}^Q_{(L-1)}, \boldsymbol{U}^M_{(L)}\Big).$$

Using a standard property of conditional expectations, we have

$$r_{V|\boldsymbol{U}^Q_{(L-1)}}\left(\boldsymbol{U}^Q_{(L-1)}\right) = \mathrm{E}_{c^+_{V|\boldsymbol{U}^Q_{(L-1)}}}\left(V|\boldsymbol{U}^Q_{(L-1)}\right) = \mathrm{E}\left[\mathrm{E}_{c^+_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}}\left(V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}\right)\middle|\,\boldsymbol{U}^Q_{(L-1)}\right]$$
$$= \mathrm{E}\left[r_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}\left(\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}\right)\middle|\,\boldsymbol{U}^Q_{(L-1)}\right].$$

Thus, an alternative expression for $\rho^2_{(\boldsymbol{X}^M_{(L)}\to Y|\boldsymbol{X}^Q_{(L-1)})}$ in (12) is given by

$$\rho^2_{(\boldsymbol{X}^M_{(L)}\to Y|\boldsymbol{X}^Q_{(L-1)})} = 12\mathrm{E}\left[\left\{r_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}\left(\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}\right) - r_{V|\boldsymbol{U}^Q_{(L-1)}}\left(\boldsymbol{U}^Q_{(L-1)}\right)\right\}^2\right]$$
$$= 12\mathrm{E}\left[\mathrm{E}\left[\left\{r_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}\left(\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}\right) - r_{V|\boldsymbol{U}^Q_{(L-1)}}\left(\boldsymbol{U}^Q_{(L-1)}\right)\right\}^2\middle|\,\boldsymbol{U}^Q_{(L-1)}\right]\right]$$
$$= 12\mathrm{E}\left[\mathrm{E}\left[\left\{r_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}\left(\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}\right) - \mathrm{E}\left[r_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}\left(\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}\right)|\boldsymbol{U}^Q_{(L-1)}\right]\right\}^2\middle|\,\boldsymbol{U}^Q_{(L-1)}\right]\right]$$
$$= 12\mathrm{E}\left[\mathrm{Var}\left[r_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}\left(\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}\right)\middle|\,\boldsymbol{U}^Q_{(L-1)}\right]\right].$$

Applying the variance decomposition theorem to $\mathrm{Var}\left[r_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}\left(\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}\right)\right]$, we obtain that

$$\rho^2_{(\boldsymbol{X}\to Y)} = 12\mathrm{Var}\left[r_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}\left(\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}\right)\right]$$
$$= 12\mathrm{Var}\left[\mathrm{E}\left[r_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}\left(\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}\right)\middle|\,\boldsymbol{U}^Q_{(L-1)}\right]\right] + 12\mathrm{E}\left[\mathrm{Var}\left[r_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}\left(\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}\right)\middle|\,\boldsymbol{U}^Q_{(L-1)}\right]\right]$$
$$= 12\mathrm{Var}\left[r_{V|\boldsymbol{U}^Q_{(L-1)}}\left(\boldsymbol{U}^Q_{(L-1)}\right)\right] + 12\mathrm{E}\left[\mathrm{Var}\left[r_{V|\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}}\left(\boldsymbol{U}^Q_{(L-1)},\boldsymbol{U}^M_{(L)}\right)\middle|\,\boldsymbol{U}^Q_{(L-1)}\right]\right] = \rho^2_{(\boldsymbol{X}^Q_{(L-1)}\to Y)} + \rho^2_{(\boldsymbol{X}^M_{(L)}\to Y|\boldsymbol{X}^Q_{(L-1)})}.$$

This concludes the proof for the sequential decomposition of $\rho^2_{(\boldsymbol{X}\to Y)}$. □

Proposition 3 is a corollary of the sequential decomposition. For Proposition 3-(a), we first apply the sequential decomposition over the partition of $\boldsymbol{X}^Q_{(\ell+1)} = \boldsymbol{X}^Q_{(\ell)} \bigcup \boldsymbol{X}^M_{(\ell+1)}$ for $\ell \in \{1,\dots,L-1\}$. Then we have $\rho^2_{(\boldsymbol{X}^Q_{(\ell+1)}\to Y)} = \rho^2_{(\boldsymbol{X}^Q_{(\ell)}\to Y)} + \rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})}$ and thus $\rho^2_{(\boldsymbol{X}^Q_{(\ell)}\to Y)} \le \rho^2_{(\boldsymbol{X}^Q_{(\ell+1)}\to Y)}$ where the equality holds if and only if $\rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} = 0$, which is equivalent to $r_{V|\boldsymbol{U}^Q_{(\ell)},\boldsymbol{U}^M_{(\ell+1)}}(\boldsymbol{u}^Q_{(\ell)},\boldsymbol{u}^M_{(\ell+1)}) = r_{V|\boldsymbol{U}^Q_{(\ell)}}(\boldsymbol{u}^Q_{(\ell)})$ for any $\boldsymbol{u}^Q_{(\ell)}$ and $\boldsymbol{u}^M_{(\ell+1)}$. For Proposition 3(b), if M^* is a new partition of $\boldsymbol{X}$ such that $\boldsymbol{X}^{M^*} = \boldsymbol{X}^M_{(\ell+1)} \bigcup \boldsymbol{X}^M_{(\ell+2)}$ for $\ell \in \{1,\dots,L-2\}$, then by the sequential decomposition formula,

$$\rho^2_{(\boldsymbol{X}^Q_{(\ell+2)}\to Y)} = \rho^2_{(\boldsymbol{X}^Q_{(\ell)}\to Y)} + \rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} + \rho^2_{(\boldsymbol{X}^M_{(\ell+2)}\to Y|\boldsymbol{X}^Q_{(\ell+1)})} = \rho^2_{(\boldsymbol{X}^Q_{(\ell)}\to Y)} + \rho^2_{(\boldsymbol{X}^{M^*}\to Y|\boldsymbol{X}^Q_{(\ell)})}.$$

Note that the second equality implies $\rho^2_{(\boldsymbol{X}^{M^*}\to Y|\boldsymbol{X}^Q_{(\ell)})} = \rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} + \rho^2_{(\boldsymbol{X}^M_{(\ell+2)}\to Y|\boldsymbol{X}^Q_{(\ell+1)})}$, which gives the proof of Proposition 3(b). □

Further, by Proposition 2 in [54], $\rho^2_{(\boldsymbol{X}^Q_{(\ell)}\to Y)}$ is invariant with respect to $\widetilde{\boldsymbol{X}}$ and $\widetilde{Y}$, respectively. That is $\rho^2_{(\widetilde{\boldsymbol{X}}^Q_{(\ell)}\to Y)} = \rho^2_{(\boldsymbol{X}^Q_{(\ell)}\to\widetilde{Y})} = \rho^2_{(\boldsymbol{X}^Q_{(\ell)}\to Y)}$ for $\ell \in \{1,\dots,L\}$. Therefore, by the sequential decomposition formula in (14), we have $\rho^2_{(\boldsymbol{X}^Q_{(\ell+1)}\to Y)} = \rho^2_{(\boldsymbol{X}^Q_{\ell}\to Y)} + \rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} = \rho^2_{(\widetilde{\boldsymbol{X}}^Q_{\ell}\to Y)} + \rho^2_{(\widetilde{\boldsymbol{X}}^M_{(\ell+1)}\to Y|\widetilde{\boldsymbol{X}}^Q_{(\ell)})} = \rho^2_{(\widetilde{\boldsymbol{X}}^Q_{(\ell+1)}\to Y)}$. This implies the conditional association measure $\rho^2_{(\widetilde{\boldsymbol{X}}^M_{(\ell+1)}\to\widetilde{Y}|\widetilde{\boldsymbol{X}}^Q_{(\ell)})}$ is invariant with respect to $\widetilde{\boldsymbol{X}}$. Similarly, it can be shown that $\rho^2_{(\widetilde{\boldsymbol{X}}^M_{(\ell+1)}\to\widetilde{Y}|\widetilde{\boldsymbol{X}}^Q_{(\ell)})}$ is invariant with respect to $\widetilde{Y}$. This gives the proofs of Proposition 4 and Theorem 2(b). □

The proof of Proposition 5. (i) If Y is jointly independent with $\boldsymbol{X}$, then $p_{\boldsymbol{i},j} = p_{\boldsymbol{i}+}p_{+j}$ for all $\boldsymbol{i}$ and j. This implies $p_{\boldsymbol{i}^M_{(\ell)},j} = p_{\boldsymbol{i}^M_{(\ell)}+}p_{+j}$ for all $\boldsymbol{i}^M_{(\ell)}$ and j. By (4), we know $\rho^2_{(\boldsymbol{X}\to Y)} = 0$ and hence $\rho^2_{(\boldsymbol{X}^M_{(\ell)}\to Y)} = 0$ for $1 \le \ell \le L$.

(ii) Suppose that Y is marginally independent with $\boldsymbol{X}^Q_{(\ell*)}$ in the marginal contingency table for $({\boldsymbol{X}^Q_{(\ell*)}}^\top, Y)^\top$, then, $p_{\boldsymbol{i}^Q_{(\ell)},j} = p_{\boldsymbol{i}^Q_{(\ell)}+}p_{+j}$ for all $\boldsymbol{i}^Q_{(\ell)}$ and j. Then, by Definition 2, we have $\rho^2_{(\boldsymbol{X}^Q_{(\ell*)}\to Y)} = 0$. From the sequential decomposition of (14) in Theorem 2, we have $\rho^2_{(\boldsymbol{X}^M_{(1)}\to Y)} = 0$ and $\rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} = 0$ for $1 \le \ell \le \ell^* - 1$. And in this case, the sequential decomposition of (14) in Theorem 2 is reduced to be,

$$\rho^2_{(\boldsymbol{X}\to Y)} = \rho^2_{(\boldsymbol{X}^M_{(1)}\to Y)} + \sum_{\ell=1}^{L-1}\rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} = \sum_{\ell=\ell^*}^{L-1}\rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})}.$$

(iii) Suppose that Y and $\boldsymbol{X}^M_{(\ell^*+1)}$ are conditionally independent given $\boldsymbol{X}^Q_{(\ell*)}$, that is $p_{j|\boldsymbol{i}^Q_{(\ell*)},\boldsymbol{i}^M_{(\ell^*+1)}} = p_{j|\boldsymbol{i}^Q_{(\ell*)}}$ for all $\boldsymbol{i}^Q_{(\ell*)}$, $\boldsymbol{i}^M_{(\ell^*+1)}$ and j. Therefore, we have $r_{V|\boldsymbol{U}^Q_{(\ell*)},\boldsymbol{U}^M_{(\ell^*+1)}}(\boldsymbol{u}^Q_{(\ell*)}, \boldsymbol{u}^M_{(\ell^*+1)}) = \sum_{j=1}^J s_j p_{j|\boldsymbol{i}^Q_{(\ell*)},\boldsymbol{i}^M_{(\ell^*+1)}} = \sum_{j=1}^J s_j p_{j|\boldsymbol{i}^Q_{(\ell*)}} = r_{V|\boldsymbol{U}^Q_{(\ell*)}}(\boldsymbol{u}^Q_{(\ell*)})$ for any $\boldsymbol{u}^Q_{(\ell*)}$ and $\boldsymbol{u}^M_{(\ell^*+1)}$. By Definition 2, this implies

$$\rho^2_{(\boldsymbol{X}^M_{(\ell^*+1)}\to Y|\boldsymbol{X}^Q_{(\ell*)})} = \frac{Var\left(r_{V|\boldsymbol{U}^Q_{(\ell*)},\boldsymbol{U}^M_{(\ell^*+1)}}(\boldsymbol{U}^Q_{(\ell*)}, \boldsymbol{U}^M_{(\ell^*+1)}) - r_{V|\boldsymbol{U}^Q_{(\ell*)}}(\boldsymbol{U}^Q_{(\ell*)})\right)}{Var(V)} = 0,$$

and then the sequential decomposition of (14) in Theorem 2 becomes

$$\rho^2_{(\boldsymbol{X}\to Y)} = \rho^2_{(\boldsymbol{X}^M_{(1)}\to Y)} + \sum_{\ell=1}^{\ell^*-1}\rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} + \sum_{\ell=\ell^*+1}^{L-1}\rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} \quad \text{for} \quad 2 \le \ell^* \le L-2,$$

Note that if $\ell^* = 1$ ($\ell^* = L - 1$), the second term (the third term) in the above equation does not exist. □

The proof of Proposition 6. (i) By the monotone property of the marginal CCRAM in Proposition 3, $\rho^2_{(\boldsymbol{X}\to Y)} = 0$ implies $\rho^2_{(\boldsymbol{X}^Q_{(\ell)}\to Y)} = 0$ for $\ell \in \{1, \dots, L\}$. Then by the sequential decomposition of CCRAM in Theorem 2, we have $\rho^2_{(\boldsymbol{X}^Q_{(\ell+1)}\to Y)} = \rho^2_{(\boldsymbol{X}^Q_{(\ell)}\to Y)} + \rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})}$ and $\rho^2_{(\boldsymbol{X}^Q_{(\ell)}\to Y)} = \rho^2_{(\boldsymbol{X}^M_{(\ell)}\to Y)} + \rho^2_{(\boldsymbol{X}^Q_{(\ell-1)}\to Y|\boldsymbol{X}^M_{(\ell)})}$. This implies that for $\ell \in \{1, \dots, L-1\}$, $\rho^2_{(\boldsymbol{X}^M_{(\ell)}\to Y)} = \rho^2_{(\boldsymbol{X}^Q_{(\ell)}\to Y)} = \rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} = 0$ and then $r_{V|\boldsymbol{U}^Q_{(\ell)},\boldsymbol{U}^M_{(\ell+1)}}(\boldsymbol{u}^Q_{(\ell)}, \boldsymbol{u}^M_{(\ell+1)}) = r_{V|\boldsymbol{U}^Q_{(\ell)}}(\boldsymbol{u}^Q_{(\ell)}) = r_{V|\boldsymbol{U}^M_{(\ell)}}(\boldsymbol{u}^M_{(\ell)}) = 1/2$ by Proposition 3 (iii) in [54].

(ii) Assume $\rho^2_{(\boldsymbol{X}\to Y)} = \rho^2_{(\boldsymbol{X}^M_{(1)}\to Y)} + \sum_{\ell=1}^{\ell^*-1}\rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})}$. From the sequential decomposition of (14) in Theorem 2, we have $\rho^2_{(\boldsymbol{X}^M_{(\ell+1)}\to Y|\boldsymbol{X}^Q_{(\ell)})} = 0$ for $\ell \in \{\ell^*, \dots, L-1\}$, and by Definition 2, we have $r_{V|\boldsymbol{U}^Q_{(\ell)},\boldsymbol{U}^M_{(\ell+1)}}(\boldsymbol{u}^Q_{(\ell)}, \boldsymbol{u}^M_{(\ell+1)}) = r_{V|\boldsymbol{U}^Q_{(\ell)}}(\boldsymbol{u}^Q_{(\ell)})$ for any $\boldsymbol{u}^Q_{(\ell)}$ and $\boldsymbol{u}^M_{(\ell+1)}$ for $\ell \in \{\ell^*, \dots, L-1\}$. The same argument applies for the case of $\ell^* = 1$ where $\rho^2_{(\boldsymbol{X}\to Y)} = \rho^2_{(\boldsymbol{X}^M_{(1)}\to Y)}$. □

Appendix B. Supplementary data

The supplementary material includes the simulation studies to examine (i) the finite-sample performance of the estimators of the proposed SCCRAMs and (ii) the sizes and powers of the permutation tests associated with the SCCRAMs which were discussed in Section 6.

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmva.2023.105179.

References

[1] A. Agresti, Considerations in measuring partial association for ordinal categorical data, J. Amer. Statist. Assoc. 72 (357) (1977) 37–45.
[2] A. Agresti, Analysis of Ordinal Categorical Data, second ed., Wiley, New York, 2010.
[3] A. Agresti, Categorical Data Analysis, third ed., John Wiley & Sons, Hoboken, New Jersey, 2013.
[4] J. Aitchison, S. Silvey, The generalization of probit analysis to the case of multiple responses, Biometrika 44 (1957) 131–140.
[5] J.A. Anderson, Regression and ordered categorical variables, J. R. Stat. Soc. Ser. B Stat. Methodol. 46 (1) (1984) 1–22.
[6] Y. Benjamini, D. Yekutieli, The control of the false discovery rate in multiple testing under dependency, Ann. Statist. 29 (2001) 1165–1188.
[7] K.J. Berry, J.E. Johnston, P.W. Mielke Jr., The Measurement of Association: A Permutation Statistical Approach, Springer, Cham, 2018.
[8] K.J. Berry, J.E. Johnston, P.W. Mielke Jr., A Primer of Permutation Statistical Methods, Springer, Cham, 2019.
[9] M. Conaway, Analysis of repeated categorical measurements with conditional likelihood methods, J. Amer. Statist. Assoc. 84 (1989) 53–62.
[10] J.A. Davis, A partial coefficient for goodman and Kruskal's gamma, J. Amer. Statist. Assoc. 62 (317) (1967) 189–193.
[11] A.C. Davison, D.V. Hinkley, Bootstrap Methods and their Application, Cambridge University Press, Cambridge, 2014.
[12] M. Denuit, P. Lambert, Constraints on concordance measures in bivariate discrete data, J. Multivariate Anal. 93 (1) (2005) 40–57.
[13] A. Dobra, K. Buhikire, J.G. Voss, Identifying mediating variables with graphical models: An application to the study of causal pathways in people living with HIV, J. Appl. Stat. 47 (7) (2020) 1298–1314.
[14] D. Donoho, 50 Years of data science, J. Comput. Graph. Statist. 26 (2017) 745–766.

[15] R. Douglas, S. Fienberg, M.-L. Lee, A. Sampson, L. Whitaker, Positive dependence concepts for ordinal contingency tables, in: H.W. Block, A.R. Sampson, T.H. Savits (Eds.), Topics in Statistical Dependence, Institute of Mathematical Statistics, Hayward, CA, 1991, pp. 189–202.
[16] K.R. E., Peripheral fatigue: High-energy phosphates and hydrogen ions, PM & R 2 (5) (2010) 347–358.
[17] B. Efron, R. Tibshirani, An Introduction to the Bootstrap, Chapman & Hall, New York, 1993.
[18] V. Farewell, A note on regression analysis of ordinal data with variability of classification, Biometrika 69 (1982) 533–538.
[19] S. Fienberg, E. Bromet, D. Follmann, D. Lambert, S. May, Longitudinal analysis of categorical epidemiological data: A study of three mile Island, Environ. Health Perspect. 63 (1985) 241–248.
[20] J. Finsterer, S. Mahjoub, Fatigue in healthy and diseased individuals, Am. J. Hospice Palliat. Care 31 (5) (2014) 562–575.
[21] A.S. Fullerton, J. Xu, Ordered Regression Models: Parallel, Partial, and Non-Parallel Alternatives, Chapman and Hall/CRC, New York, 2016.
[22] C. Genest, J.G. Nešlehová, B. Rémillard, On the empirical multilinear copula process for count data, Bernoulli 20 (3) (2014) 1344–1371.
[23] C. Genest, J.G. Nešlehová, B. Rémillard, Asymptotic behavior of the empirical multilinear copula process under broad conditions, J. Multivariate Anal. 159 (2017) 82–110.
[24] P. Good, Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses, Springer, New York, 2000.
[25] L. Goodman, Simple models for the analysis of association in cross-classifications having ordered categories, J. Amer. Statist. Assoc. 74 (1979) 537–552.
[26] L. Goodman, The analysis of dependence in cross-classifications having ordered categories, using log-linear models for frequencies and log-linear models for odds, Biometrics 39 (1983) 149–160.
[27] L. Goodman, The analysis of cross-classified data having ordered and/or unordered categories: Association models, correlation models, and asymmetry models for contingency tables with or without missing entries, Ann. Statist. 13 (1985) 10–69.
[28] L.A. Goodman, W.H. Kruskal, Measures of association for cross classifications, J. Amer. Statist. Assoc. 49 (268) (1954) 732–764.
[29] S. Guo, E. Thompson, Analysis of sparse contingency tables: Monte Carlo estimation of exact p values, Department of Statistics, University of Washington, 1989.
[30] F.E. Harrell, Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis, Springer, New York, 2015.
[31] H. Joe, Dependence Modeling with Copulas, Chapman and Hall/CRC, New York, 2014.
[32] M.G. Kendall, A new measure of rank correlation, Biometrika 30 (1/2) (1938) 81–93.
[33] M. Kendall, Rank Correlation Methods, C. Griffin, London, 1948.
[34] M.G. Kendall, A. Stuart, The advanced theory of statistics, vol. 2: Inference and Relationship, Griffin, London, 1961.
[35] V. Kuzmenko, R. Salam, S. Uryasev, Checkerboard copula defined by sums of random variables, Depend. Model. 8 (1) (2020) 70–92.
[36] E. Läärä, J. Matthews, The equivalence of two models for ordinal data, Biometrika 72 (1985) 206–207.
[37] J. Landis, E. Heyman, G. Koch, Average partial association in three-way contingency tables: A review and discussion of alternative tests, Internat. Statist. Rev. 46 (1978) 237–254.
[38] I. Liu, A. Agresti, The analysis of ordered categorical data: An overview and a survey of recent developments, Test 14 (1) (2005) 1–73.
[39] Q. Liu, C. Li, V. Wanga, B.E. Shepherd, Covariate-adjusted Spearman's rank correlation with probability-scale residuals, Biometrics 74 (2) (2018) 595–605.
[40] N. Mantel, Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure, J. Amer. Statist. Assoc. 58 (1963) 690–700.
[41] P. McCullagh, Regression models for ordinal data, J. R. Stat. Soc. Ser. B 42 (1980) 109–142.
[42] R.B. Nelsen, An Introduction to Copulas, Springer, New York, 2006.
[43] J. Nešlehová, On rank correlation measures for non-continuous random variables, J. Multivariate Anal. 98 (3) (2007) 544–567.
[44] J. Rachasingho, S. Tasena, A metric space of subcopulas—An approach via Hausdorff distance, Fuzzy Sets and Systems 378 (2020) 144–156.
[45] L. Rüschendorf, On the distributional transform, Sklar's theorem, and the empirical copula process, J. Statist. Plann. Inference 139 (2009) 3921–3927.
[46] B. Schweizer, A. Sklar, Operations on distribution functions not derivable from operations on random variables, Studia Math. 52 (1) (1974) 43–52.
[47] G. Shmueli, To explain or to predict? Statist. Sci. 25 (3) (2010) 289–310.
[48] G. Simon, Alternative analyses for the singly-ordered contingency table, J. Amer. Statist. Assoc. 69 (1974) 971–976.
[49] A. Sklar, Fonctions de répartition à n dimensions et leurs marges, Publicat. de L'Institut de Stat. de L'Univ. de Paris 8 (1959) 229–231.
[50] M. Smithson, E.C. Merkle, Generalized Linear Models for Categorical and Continuous Limited Dependent Variables, CRC Press, Boca Raton, FL, 2013.
[51] B. Sutradhar, Longitudinal Categorical Data Analysis, Springer, New York, 2014.
[52] J. Tukey, Exploratory Data Analysis, Addison Wesley, Reading, MA, 1977.
[53] G. Tutz, Regression for Categorical Data, Cambridge University Press, Cambridge, 2011.
[54] Z. Wei, D. Kim, On exploratory analytic method for multi-way contingency tables with an ordinal response variable and categorical explanatory variables, J. Multivariate Anal. 186 (2021) 104793.