

Modelling contingency tables with hierarchical logits and odds ratios ^(*)

Modelli per tabelle di contingenza con logit e odds ratio gerarchici

Manuela Cazzaro¹ and Roberto Colombi²

¹ Dipartimento di Metodi Quantitativi per le Scienze Economiche ed Aziendali,
Università di Milano-Bicocca, e-mail: manuela.cazzaro@unimib.it

² Dipartimento di Ingegneria Gestionale e dell'Informazione, Università di Bergamo

Riassunto: Per modellare un'ipotesi di dipendenza monotona tra due variabili categoriali ordinali $\mathcal{A}$ e $\mathcal{B}$, cioè il fatto che $\mathcal{A}$ tende ad essere stocasticamente più grande o più piccola al crescere di $\mathcal{B}$, in letteratura sono di solito utilizzati gli ordinamenti stocastici semplice, uniforme e del rapporto delle verosimiglianze. Ognuna delle ipotesi di dipendenza monotona appena citate è equivalente alla non-negatività degli odds ratios generalizzati di una opportuna famiglia. In questo lavoro viene introdotto un nuovo ordinamento stocastico, denominato gerarchico, e viene dimostrato il suo legame con la famiglia degli odds ratios gerarchici.

Keywords: stochastic orderings, contingency tables, hierarchical odds ratios, order restricted inference.

1. Introduction

Let $\mathcal{A}$ and $\mathcal{B}$ be two ordered categorical variables with r and c categories respectively. In the contingency table defined by the joint distribution of the two variables an hypothesis of monotone dependence of $\mathcal{A}$ from $\mathcal{B}$ states that the conditional distributions of $\mathcal{A}$ can be ordered according to a stochastic dominance criterion in a coherent way with the order of the categories of the variable $\mathcal{B}$. In order to formalize the previous notion of monotone dependence, the stochastic simple order, the uniform stochastic order and the likelihood ratio order are usually introduced. These orderings correspond to increasingly restrictive monotone dependence concepts (Dardanoni and Forcina, 1998). Every previous monotone dependence hypothesis is equivalent to the hypothesis that the logarithms of $(r-1)(c-1)$ generalized odds ratios (o.r.) of a convenient type are non-negative. The type of generalized odds ratios depends on the stochastic order which is used. In this paper a new ordering, called hierarchical stochastic order, is introduced and it is shown that the monotone dependence hypothesis based on it is equivalent to the hypothesis that the logarithms of $(r-1)(c-1)$ hierarchical odds ratios are non-negative.

2. Hierarchical logits and odds ratios

Let X be an ordered categorical variable with categories in the set $A = \{a_i : i = 1, \dots, r\}$ and let $G(A)$ be a *hierarchy of sets*, or *hierarchical family of sets* defined on A . By *hierarchy of sets*, defined on $A = \{a_i : i = 1, \dots, r\}$, a family $G(A)$ of subsets of A is

(*) Work partially supported by a MIUR2004 grant.

intended such that: $\{a_i\} \in G(A), i = 1, \dots, r; A \in G(A)$; two sets of $G(A)$ are disjoint or one is a subset of the other; each non-minimal set, according to the inclusion order, N_i of $G(A)$, is the union of two disjoint and non-empty sets belonging to $G(A)$, that is $N_i = P_i \cup Q_i, P_i \in G(A), Q_i \in G(A)$.

It can be shown that if A has r elements then the non-minimal sets of $G(A)$ are $r - 1$.

A *hierarchy family of sets* is said to be *coherent* according to the order of the categories $a_i \in A$ of a ordered categorical variable if $N_i = P_i \cup Q_i$ implies that $\max\{a_i : a_i \in P_i\}$ and $\min\{a_i : a_i \in Q_i\}$ are adjacent categories. From this it follows, by induction, that the sets of a *coherent hierarchy of sets* $G(A)$ contain contiguous categories.

For each non-minimal set $N_i = P_i \cup Q_i, i = 1, \dots, r - 1$, of $G(A)$ the hierarchical logit $\ln \frac{Pr(Q_i)}{Pr(P_i)}$ is defined. Furthermore given two random variables (r.v.) X, Y , with the same set of categories A and *coherent hierarchy of sets* $G(A)$, the odds ratios $\frac{Pr(X \in P_i)}{Pr(X \in Q_i)} \cdot \frac{Pr(Y \in Q_i)}{Pr(Y \in P_i)}, i = 1, \dots, r - 1$, are called *hierarchical odds ratios*.

3. Hierarchical order and relationships with the simple stochastic order

With the notation $[X|X \in N]$ is intended the r.v. X , truncated on the set N , with probability function $\frac{Pr(X=a_i)}{Pr(X \in N)}, \forall a_i \in N$. In order to define the monotone dependence hypotheses we introduce the following stochastic orderings.

Likelihood Ratio Order (LR): if $[X|a \leq X \leq b]$ is stochastically smaller than $[Y|a \leq Y \leq b]$ with respect to the simple stochastic order for all $a_1 < a < b \leq a_r$ then the r.v. X is dominated by the r.v. Y with respect to the likelihood ratio ordering, $(X \preceq_{LR} Y)$.

Uniform Order (U): if $[X|X \geq a]$ is stochastically smaller than $[Y|Y \geq a]$ with respect to the simple stochastic order for all $a_1 < a \leq a_r$ then the r.v. X is dominated by the r.v. Y with respect to the uniform ordering, $(X \preceq_U Y)$.

Hierarchical Order ($H_{G(A)}$): if $[X|X \in N_i]$ is stochastically smaller than $[Y|Y \in N_i]$ with respect to the simple stochastic order for all non-minimal sets N_i belonging to the coherent hierarchical family $G(A)$ then the r.v. X is dominated by the r.v. Y with respect to the hierarchical ordering, $(X \preceq_{H_{G(A)}} Y)$.

It is worthwhile to note that the uniform ordering is a particular case of hierarchical ordering defined by the coherent hierarchical family $G(A)$ with non-minimal sets $N_i = \{a_i, a_{i+1}, \dots, a_r\}, i = 1, 2, \dots, r - 1$.

Moreover from the previous definitions it follows directly that the $LR, U, H_{G(A)}$ orders imply the simple stochastic order, S , and are implied by the likelihood ratio order LR .

The following theorem shows that every previous ordering is equivalent to the non-negativity of the logarithms of $(r - 1)(c - 1)$ generalized odds ratios (o.r.).

Theorem.

i) the LR order is equivalent to the non-negativity hypothesis of the $(r - 1)$ log-local-o.r.: $\ln \frac{Pr(X=a_i)Pr(Y=a_{i+1})}{Pr(X=a_{i+1})Pr(Y=a_i)}, i = 1, \dots, r - 1$;

ii) the U order is equivalent to the non-negativity hypothesis of the $(r - 1)$ log-continuation o.r.: $\ln \frac{Pr(X=a_i) \cdot \sum_{i=1}^{r-1} Pr(Y=a_{i+1})}{Pr(Y=a_i) \cdot \sum_{i=1}^{r-1} Pr(X=a_{i+1})}, i = 1, \dots, r - 1$;

iii) the $H_{G(A)}$ order is equivalent to the non-negativity hypothesis of the logarithms of the $(r - 1)$ hierarchical o.r.

Proof. As the first two statements regarding the LR and U orders are well-known

(Müller and Stoyan, 2002, and Dardanoni and Forcina, 1998) we only give the proof of the third point. In the following proof "stochastically smaller with respect to the simple stochastic order" will be abbreviated by "stochastically smaller". Let us assume that the hierarchical o.r. are greater or equal than one. If $N_i \in G(A)$ has cardinality $k = 2$ the thesis is trivially true. Let N_i be of cardinality $k > 2$ and let us assume that the thesis is true for the non-minimal sets of $G(A)$ of cardinality smaller than k . By definition $N_i = P_i \cup Q_i$ where P_i and Q_i have obviously cardinality smaller than k . By the induction hypothesis $[X|X \in P_i]$ ($[X|X \in Q_i]$) is stochastically smaller than $[Y|Y \in P_i]$ ($[Y|Y \in Q_i]$). It follows that a mixture of the random variables $[X|X \in P_i]$ and $[X|X \in Q_i]$ defines a r.v. that is stochastically smaller than the r.v. given by the mixture of $[Y|Y \in P_i]$ and $[Y|Y \in Q_i]$ whatever the weights of the mixture might be. In particular, this is true if the weights $Pr(X \in P_i|N_i)$ and $Pr(X \in Q_i|N_i)$ are utilised; these weights if applied to $[X|X \in P_i]$ and $[X|X \in Q_i]$ give the random variable $[X|X \in N_i]$. If the mixture of $[Y|Y \in P_i]$ and $[Y|Y \in Q_i]$ is considered with weights $Pr(X \in P_i|N_i)$ and $Pr(X \in Q_i|N_i)$ a stochastically smaller r.v. is obtained compared to the one obtained considering the weights $Pr(Y \in P_i|N_i)$ and $Pr(Y \in Q_i|N_i)$ as $\frac{Pr(X \in P_i)Pr(Y \in Q_i)}{Pr(X \in Q_i)Pr(Y \in P_i)} \geq 1$. Since in this last case the random variable $[Y|Y \in N_i]$ is obtained, the statement follows by induction. That the hierarchical order implies the non-negativity of the log-hierarchical o.r. follows directly from the definition of the stochastic order.

4. Hierarchical monotone dependence hypothesis

Let us consider a contingency table defined by two ordered categorical variables: the row variable $\mathcal{A}$ and the column variable $\mathcal{B}$ with categories $a_1, \dots, a_r$ and $b_1, \dots, b_c$ respectively. An hypothesis of monotone dependence of $\mathcal{A}$ from $\mathcal{B}$ states that the conditional distributions of $\mathcal{A}$ can be ordered according to a stochastic dominance criterion in a coherent way with the order of the categories of the variable $\mathcal{B}$. From now on the focus will be on the new $H_{G(A)}$ order introduced in the previous section; for details on the other orderings see Dardanoni and Forcina (1998). When the dominance criterium is given by a Hierarchical Order $H_{G(A)}$ the dependence hypothesis is called hierarchical monotone dependence hypothesis.

If the $H_{G(A)}$ order is adopted then the equivalent condition of non-negativity of the hierarchical odds ratios, given in the theorem of the previous section, has to be satisfied for each couple of adjacent columns of the contingency table. Let us note that for each couple of adjacent columns it is possible to calculate $(r - 1)$ hierarchical odds ratios. Considering $(c - 1)$ couples of adjacent columns $(r - 1)(c - 1)$ odds ratios are obtained. Thus the hierarchical monotone dependence hypothesis is equivalent to the non-negativity of $(r - 1)(c - 1)$ hierarchical odds ratios.

If θ is the parameters vector of a saturated log-linear model then the joint probabilities are the elements of the vector $\pi(\theta)$ defined as follows (Colombi and Forcina, 2001): $\ln(\pi(\theta)) = \mathbf{Z}\theta - \mathbf{1} \ln[\mathbf{1}' \exp(\mathbf{Z}\theta)]$, where $\mathbf{Z}$ is a design matrix such that $[\mathbf{Z}, \mathbf{1}]$ is non-singular. The vector of the $(r - 1)(c - 1)$ non-negativity constraints on the hierarchical odds ratios can be written as follows: $\mathbf{C} \ln \mathbf{M}\pi(\theta) \geq 0$ where $\mathbf{C}$ is a matrix of linear independent contrasts. The matrix $\mathbf{M}$, the elements of which are equal to 0 or 1, depends on the hierarchical family of sets $G(A)$. This way of building the inequality constraints of the $H_{G(A)}$ hypothesis is analogous to that utilised by Colombi and Forcina, (2001), in

the LR , U and S order cases. It is worthwhile to note that the constraints on the o.r. are non-linear constraints on the log-linear parameter of the model. It is possible to show that the jacobian matrix of the constraints has full row rank (an extended version of this paper with the proof is available from the authors); thus the asymptotic chi-bar distribution of the likelihood ratio statistics to test the hypothesis $H_{G(A)}$ against the unrestricted alternative hypothesis follows (cfr. Silvapulle and Sen, 2005). An analogous result holds if the hypothesis of stochastic independence is tested against the hierarchical monotone dependence hypothesis (cfr. Colombi and Forcina, 2001).

5. Application

The models proposed in this work are used to analyze the data of Table 1, where medical service's users are asked to evaluate their satisfaction (*unsatisfied* (U), *satisfied* (S), *really satisfied* (RS), *extremely satisfied* (ES)) regarding the waiting-room's comfort (*COMFORT*) and the satisfaction (with the same categories U, S, RS, ES) regarding the perception of the waiting-time (*TIME*) before a specialist examination is carried out. The uniform dependence of the perception of the waiting-time from the comfort of the waiting-room is tested. Note that the coherent hierarchy of sets considered is the one with non-minimal sets: $N_1 = \{U, S\}$, $N_2 = \{RS, ES\}$, $N_3 = \{U, S, RS, ES\}$.

Table 1: *The data.*

TIME / COMFORT	U	S	RS	ES
U	86	31	50	39
S	34	40	42	34
RS	53	50	99	76
ES	43	44	121	367

Testing the different models against the unrestricted alternative hypothesis on these data gives the following results: i) *likelihood ratio monotone dependence* model: $G^2 = 2.22$, $p\text{-value} = 0.9232$; ii) *uniform monotone dependence* model: $G^2 = 2.13$, $p\text{-value} = 0.8644$; iii) *hierachical monotone dependence* model: $G^2 = 0.56$, $p\text{-value} = 0.9859$; iv) *simple monotone dependence* model: $G^2 = 0.00$, $p\text{-value} = 0.9990$. All the models show an excellent fit and of course the ordering of the values of the G^2 statistics is coherent with the implications between the monotone dependence hypotheses.

References

- Colombi R., Forcina A. (2001) Marginal regression models for the analysis of positive association of ordinal response variables *Biometrika*, 88, 1007–1019.
- Dardanoni V., Forcina, A. (1998) A unified approach to likelihood inference on stochastic orderings in a nonparametric context, *Journal of the American Statistical Association*, 93, 1112–1123.
- Müller A., Stoyan D. (2002) *Comparison methods for stochastic models and risks*, Wiley, New York.
- Silvapulle M.J., Sen P.K. (2005) *Constrained statistical inference*, Wiley, New Jersey.