Hierarchical Multinomial Marginal Models and the R-package hmmm: a brief introduction

Manuela Cazzaro[1], Roberto Colombi[2]

[1] Dip. Met. Quant. per le Sc. Ec. Az., Università di Milano Bicocca - Piazza dell'Ateneo Nuovo, 1 - 20126 Milano - Italy; e-mail: `manuela.cazzaro@unimib.it`

[2] Dip. Ingegneria dell'Informazione e Metodi Matematici, Università di Bergamo - Viale Marconi, 5 - 24044 Dalmine - Italy; e-mail: `colombi@unibg.it`

Abstract *Hierarchical Multinomial Marginal Models (HMMM) are introduced and the use of the package hmmm in order to specify, estimate and testing HMMM models is briefly explained. This paper is intended to be an introduction both to the theory of HMMM models and to the main features of the package hmmm.*

1 Introduction

Hierarchical Multinomial Marginal Models, (HMMM), allows logits of local, baseline, global or recursive type to be considered for each variable of a multi-way contingency table. Furthermore higher dimension interactions can be defined in correspondence to different marginal probability functions coherently with the type of logits associated to each variable. The specification of a *HMMM* depends on the definitions of *complete and hierarchical family of interaction sets* and of *generalized marginal interactions.* Bartolucci, Colombi and Forcina (2007) showed that the generalized marginal interactions associated to a complete and hierarchical family of interaction sets are a parameterization of the joint probabilities of a multi-way contingency table. *HMMM* models are special cases of Lang's (2005) *HLP* models and the Lang's asymptotic results apply also to this context. The *hmmm* R-package, that has been developed by the authors of this paper, can be used to specify and estimate *HMMM* models under equality and inequality constraints. The package is based on an extension of the Lang (2005) procedure *mph.fit*. The main extensions are about the specification of very general *HMMM* models in a friendly way and about the inference on parameters under inequality constraints. A procedure to simulate the *p-values* of the chi-bar distribution, needed to test inequalities, is also implemented.

The package *hmmm* can also be used to estimate the more general models *MPH* and *HLP* of Lang (2004, 2005) under inequality constraints.

HMMM models are defined in section two. In section three the use of the package *hmmm* is introduced and some simple examples of *HMMM* models are given. Section four describes *HMMM* models and the use of the package *hmmm* when covariates are present.

2 Multi-way contingency tables

We now consider the joint probability function of q ordinal categorical variables $A_1, \ldots, A_q$, where A_j has categories in $\mathcal{A}_j = \{a_{ji_j}, i_j = 1, 2, \ldots, r_j\}$. For ordinal variables the numbering of the categories is assumed to be coherent with their order. The set of variables that defines a given marginal distribution will be denoted by the set $\mathcal{M}$ of indices of the corresponding variables. $\mathcal{M}$ will be called marginal set and the distribution associated with it will be called $\mathcal{M}$-marginal distribution. $\mathcal{Q} = \{1, \ldots, q\}$ will refer to the joint distribution and the vector of the $r = \prod_1^q r_j$ joint probabilities will be denoted by $\boldsymbol{\pi}$. Given a vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_q)'$ of q components $\boldsymbol{x}_\mathcal{M}$ will denote the vector with components $x_j : j \in \mathcal{M}$. $\mathbf{1}$ will denote a vector of ones when the dimension is clear from the context and $\mathbf{1}_\mathcal{M}$ a vector of ones of dimension equal to the cardinality of $\mathcal{M}$.

2.1 Generalized marginal interactions

We now extend the Bartolucci, Colombi and Forcina (2007) definition of interaction parameters in order to include recursive or nested logits and log-odds ratios (o.r.), (Cazzaro and Colombi, 2006b, 2007) together with the well known types of logits: local, baseline, global, continuation and reverse-continuation and the types of log-odds ratios discussed by Douglas *et al.* (1990).

Given $r_j - 1$ pairs $\mathcal{B}_j(m_j, 0)$, $\mathcal{B}_j(m_j, 1)$, $m_j = 1, 2, \ldots, r_j - 1$, of disjoint subsets of $\mathcal{A}_j$, the logits, defined on a marginal distribution, are the log-probability odds: $\ln \frac{P(A_j \in B_j(m_j, 1))}{P(A_j \in B_j(m_j, 0))}$.

The sets $\mathcal{B}_j(m_j, 0)$ are equal to $\{a_{jm_j}\}$ for local logits and to $\{a_{ji_j} : i_j = 1, \ldots, m_j\}$ for global logits, $m_j = 1, 2, \ldots, r_j - 1$; similarly, the sets $\mathcal{B}_j(m_j, 1)$ are equal to $\{a_{j(m_j+1)}\}$ for local logits and to $\{a_{ji_j} : i_j = m_j + 1, \ldots, r_j\}$ for global logits, $m_j = 1, 2, \ldots, r_j - 1$. To consider explicitly baseline logits define the sets $\mathcal{B}_j(m_j, 0)$ to be equal to $\{a_{j1}\}$ and the sets $\mathcal{B}_j(m_j, 1)$ to be equal to $\{a_{j(m_j+1)}\}$ for any $m_j < r_j$.

For recursive or nested logits the sets $\mathcal{B}_j(m_j, 0) = \mathcal{B}_j(a_{ji_j(m_j)}, 0)$ and $\mathcal{B}_j(m_j, 1) = \mathcal{B}_j(a_{ji_j(m_j)}, 1)$, of categories of a *Coherent Complete Hierarchy of Sets* are defined as described in the Appendix for any $m_j < r_j$.

Continuation and reverse-continuation logits are special cases of recursive or nested logits (Cazzaro and Colombi, 2007).

Given a logit type for the categorical variable A_1 and a logit type for the categorical variable A_2 a family of generalized o.r., defined on a bivariate distribution, is composed by the standard odds ratios computed in the following $(r_1 - 1)(r_2 - 1)$ two by two tables T_{m_1,m_2}, $m_1 = 1, 2, ..., r_1 - 1$, $m_2 = 1, 2, ..., r_2 - 1$:

Table 1 2×2 table where a generalized o.r. is defined.

T_{m_1,m_2}	0	1
0	$pr(A_1 \in \mathcal{B}_1(m_1,0), A_2 \in \mathcal{B}_2(m_2,0))$	$pr(A_1 \in \mathcal{B}_1(m_1,0), A_2 \in B_2(m_2,1))$
1	$pr(A_1 \in \mathcal{B}_1(m_1,1), A_2 \in B_2(m_2,0))$	$pr(A_1 \in \mathcal{B}_1(m_1,1), A_2 \in B_2(m_2,1))$

When the same logit type is used for A_1 and for A_2 a family of symmetric odds ratios is defined, otherwise a family of asymmetric odds ratios is defined. A family of odds ratios is denoted by the name of the logit type used for A_1 and by the name of the logit type used for A_2 (local-global o.r., local-continuation o.r., global-continuation o.r., global-local o.r., continuation-local o.r., continuation-global o.r., etc.). If the same logit type is used for both variables the name is not repeated (local o.r., global o.r., continuation o.r., etc.). Recursive and local-recursive odds ratios have been introduced by Cazzaro and Colombi (2006b, 2007) who used them to define new types of monotone dependence hypotheses. The logits defined on marginal distributions and the odds ratios defined on bivariate distributions, just introduced, are special cases of the *generalized marginal interactions* introduced by Bartolucci, Colombi and Forcina (2007).

In order to introduce the *generalized marginal interactions* we define the marginal probabilities:

$$p_{\mathcal{M}}(\boldsymbol{m}_{\mathcal{M}}; \boldsymbol{h}_{\mathcal{M}}) = P(A_j \in \mathcal{B}_j(m_j, h_j), \ \forall j \in \mathcal{M}),$$

where $\boldsymbol{m}_{\mathcal{M}}$ is a row vector of integers m_j, $m_j < r_j$, $j \in \mathcal{M}$, and $\boldsymbol{h}_{\mathcal{M}}$ is a row vector whose elements, h_j, $j \in \mathcal{M}$, are equal to zero or to one. These marginal probabilities are probabilities of a table where the variables A_j, $\forall j \in \mathcal{M}$, have been dichotomized according to the categories $\mathcal{B}_j(m_j, 0)$ and $\mathcal{B}_j(m_j, 1)$ and where the variables A_j, $\forall j \notin \mathcal{M}$, have been marginalized. Note that different $\boldsymbol{m}_{\mathcal{M}}$ denote different tables while different $\boldsymbol{h}_{\mathcal{M}}$ denote different probabilities within the same table. In general $\mathcal{B}_j(m_j, 0) \cup \mathcal{B}_j(m_j, 1) \subset \mathcal{A}_j$, thus the probabilities of the previous tables do not always sum to one. This aspect is irrelevant to what follows.

The *generalized marginal interactions* $\eta_{\mathcal{I};\mathcal{M}}(\boldsymbol{m}_{\mathcal{I}})$ are standard baseline log-linear interactions defined in the previous marginalized and aggregated tables. Any *generalized marginal interaction* is defined by the *interaction set* $\mathcal{I}$ of the variables involved, by the $\mathcal{M}$-*marginal distribution* where it is defined and by the logit type assigned to each variable of $\mathcal{M}$. Moreover every interaction is computed in the marginalized-aggregated table where the variables A_j, $\forall j \in \mathcal{M} \backslash \mathcal{I}$, which are not involved in the interaction,

belongs to $\mathcal{B}_j(1,0) \cup \mathcal{B}_j(1,1)$. This means that the conditioning variables A_j, $\forall j \in \mathcal{M} \setminus \mathcal{I}$ are fixed at the reference set $\mathcal{B}_j(1,0)$ which is the singleton $\{a_{j1}\}$ in the case of global, local, baseline and continuation logits.

A more formal definition of the *generalized marginal interactions* is:

$$\eta_{\mathcal{I};\mathcal{M}}(\boldsymbol{m}_{\mathcal{I}}) = \sum_{\mathcal{K} \subseteq \mathcal{I}} (-1)^{|\mathcal{I} \setminus \mathcal{K}|} \log p_{\mathcal{M}}(\boldsymbol{m}_{\mathcal{I}}, \boldsymbol{1}_{\mathcal{M} \setminus \mathcal{I}}; \boldsymbol{0}_{\mathcal{M} \setminus \mathcal{K}}, \boldsymbol{1}_{\mathcal{K}}). \qquad (1)$$

The choice of the reference category $\mathcal{B}_j(1,0)$ for the conditioning variables belonging to $\mathcal{M} \setminus \mathcal{I}$ is convenient but may be arbitrary. For a discussion of the relations between generalized marginal interactions that treat the variables in $\mathcal{M} \setminus \mathcal{I}$ in a different way see Bartolucci, Colombi and Forcina (2007, sections 2.1, 2.2) who showed that only when all the variables in $\mathcal{M} \setminus \mathcal{I}$ have logits of local (baseline) type, the parameters defined by fixing the conditioning variables to a different reference category are linearly related to the ones defined by (1). It follows that only when all the variables in $\mathcal{M} \setminus \mathcal{I}$ have logits of local (baseline) type, linear constraints on the interactions (1) may be written as linear constraints on parameters defined by choosing a different reference category.

Note that the kind of dichotomy implied by the type of logits adopted for each variable should carry over when defining higher order interactions within the same marginal distribution, but not necessarily between different marginal distributions.

To exemplify the previous definitions and the associated notation consider the marginal set $\mathcal{M} = \{1,2,3,4\}$ where global logits are assigned to the first two variables and local logits are assigned to the last two variables. If A_1 and A_2 have three categories the interactions $\boldsymbol{\eta}_{\{1,2\};\{1,2,3,4\}}(\boldsymbol{m}_{\{1,2\}})$ are the four log-global logits that are standard log-odds ratios computed in the following 2×2 tables (for simplicity we omit the reference to the marginal set within the table):

Table 2 Tables where the interactions $\boldsymbol{\eta}_{\{1,2\};\{1,2,3,4\}}(\boldsymbol{m}_{\{1,2\}})$ are defined.

$A_3 = a_{31}, A_4 = a_{41}$

		A_1				A_1	
		a_{11}	a_{12}, a_{13}			a_{11}, a_{12}	a_{13}
A_2	a_{21}	$p(1111;0000)$	$p(1111;1000)$	a_{21}		$p(2111;0000)$	$p(2111;1000)$
	a_{22}, a_{23}	$p(1111;0100)$	$p(1111;1100)$	a_{22}, a_{23}		$p(2111;0100)$	$p(2111;1100)$

		A_1				A_1	
		a_{11}	a_{12}, a_{13}			a_{11}, a_{12}	a_{13}
A_2	a_{21}, a_{22}	$p(1211;0000)$	$p(1211;1000)$	a_{21}, a_{22}		$p(2211;0000)$	$p(2211;1000)$
	a_{23}	$p(1211;0100)$	$p(1211;1100)$	a_{23}		$p(2211;0100)$	$p(2211;1100)$

In the tables of the first column can also be defined the two log-global logits: $\boldsymbol{\eta}_{\{2\};\{1,2,3,4\}}(\boldsymbol{m}_{\{2\}})$ and in the tables of the first row the two log-global logits $\boldsymbol{\eta}_{\{1\};\{1,2,3,4\}}(\boldsymbol{m}_{\{1\}})$ can be defined.

If A_3, A_4 are dicothomous the third order interactions $\boldsymbol{\eta}_{\{1,2,3\};\{1,2,3,4\}}(\boldsymbol{m}_{\{1,2,3\}})$ are differences between the logs of the global o.r. of the previous tables and the logs of the global odds ratios computed in the corresponding table where $A_3 = a_{32}, A_4 = a_{41}$. These tables are reported below.

Table 3 Tables needed for the interactions $\boldsymbol{\eta}_{\{1,2,3\};\{1,2,3,4\}}(\boldsymbol{m}_{\{1,2,3\}})$.

$$A_3 = a_{32}, A_4 = a_{41}$$
$$A_1$$

		a_{11}	a_{12}, a_{13}		a_{11}, a_{12}	a_{13}
A_2	a_{21}	$p(1121;0000)$	$p(1121;1000)$	a_{21}	$p(2121;0000)$	$p(2121;1000)$
	a_{22}, a_{23}	$p(1121;0100)$	$p(1121;1100)$	a_{22}, a_{23}	$p(2121;0100)$	$p(2111;1100)$

$$A_1$$

		a_{11}	a_{12}, a_{13}		a_{11}, a_{12}	a_{13}
A_2	a_{21}, a_{22}	$p(1221;0000)$	$p(1221;1000)$	a_{21}, a_{22}	$p(2221;0000)$	$p(2221;1000)$
	a_{23}	$p(1221;0100)$	$p(1221;1100)$	a_{23}	$p(2221;0100)$	$p(2221;1100)$

The interactions $\boldsymbol{\eta}_{\{1,3\};\{1,2,3,4\}}(\boldsymbol{m}_{\{1,3\}})$ are two log-global-local odds ratios that are the standard log-odds ratios computed using the four probabilities which are located in the first row of the two N-E tables (resp. N-W) of Table 2 and Table 3.

The interactions $\boldsymbol{\eta}_{\{2,3\};\{1,2,3,4\}}(\boldsymbol{m}_{\{2,3\}})$ are two log-global-local odds ratios that are the standard log-odds ratios computed using the four probabilities which are located in the first column of the two N-E tables (resp. S-E) of Table 2 and Table 3.

Similar remarks hold for the interactions:

$$\boldsymbol{\eta}_{\{1,4\};\{1,2,3,4\}}(\boldsymbol{m}_{\{1,4\}}), \boldsymbol{\eta}_{\{2,4\};\{1,2,3,4\}}(\boldsymbol{m}_{\{2,4\}}), \boldsymbol{\eta}_{\{1,2,4\};\{1,2,3,4\}}(\boldsymbol{m}_{\{1,2,4\}}).$$

To compute the previous interactions and the fourth order interactions $\boldsymbol{\eta}_{\{1,2,3,4\};\{1,2,3,4\}}(\boldsymbol{m}_{\{1,2,3,4\}})$ which are contrasts of four log-global odds ratios, also the eight tables where $A_4 = a_{42}$ are needed. We leave to the reader the task to find out that interactions $\boldsymbol{\eta}_{\{1,3,4\};\{1,2,3,4\}}(\boldsymbol{m}_{\{1,3,4\}})$ and $\boldsymbol{\eta}_{\{2,3,4\};\{1,2,3,4\}}(\boldsymbol{m}_{\{2,3,4\}})$ are differences between the logarithms of two global-local odds ratios. The remaining three families of interactions involve only A_3 and A_4 and a sensible choice is to define them in the marginal distribution of these two variables when we are interested in the dependence of A_1 and A_2 on A_3 and A_4.

The previous example shows two important facts. First of all *generalized marginal interactions* are standard log-linear interactions which are computed in tables obtained by marginalizing with respect to some variables and by aggregating the categories of some other variables. Secondly every *generalized marginal interaction* can be seen as a contrast of well known types of generalized logits and odds ratios.

Glonek and McCullagh (1995), Colombi and Forcina (2001) showed that a vector of these interactions has the matricial representation $\boldsymbol{C}\ln(\boldsymbol{M}\boldsymbol{\pi})$. In the previous formula $\boldsymbol{\pi}$ is the vector of the joint probabilities, $\boldsymbol{C}$ is a matrix

of contrasts and $\boldsymbol{M}$ is a matrix of zeros and ones. Cazzaro and Colombi (2007) extended the previous result in order to include recursive logits and odds ratios.

2.2 Generalized marginal interactions associated to complete and hierarchical families of interaction sets

We now examine the problem of allocating the interaction sets $\mathcal{I}$ among the marginal sets within which they may be defined. Remember that a marginal distribution is denoted by the set $\mathcal{M} \subset \mathcal{Q}$ of indices of the corresponding variables. Denote by $\mathcal{F}_k$ the family of interaction sets defined within the $\mathcal{M}_k$-marginal distribution. Let also $\mathcal{P}_k$ be the family of all non-empty subsets of $\mathcal{M}_k$.

Given a sequence of marginal sets $\mathcal{M}_1, \ldots, \mathcal{M}_s$, such that $\mathcal{M}_k$ is not a subset of $\mathcal{M}_h$ for every $h < k$, a family of interaction sets is called *complete* and *hierarchical* if (i) any interaction set is defined in one $\mathcal{M}_k$-marginal distribution, (ii) $\mathcal{F}_1 = \mathcal{P}_1$ and $\mathcal{F}_k = \mathcal{P}_k \backslash \bigcup_{h<k} \mathcal{F}_h$.

The previous definition implies that $\mathcal{M}_s = \mathcal{Q}$, that every interaction set belongs to only one marginal distribution, that $\mathcal{M}_k \in \mathcal{F}_k$, for every k and that if a set $\mathcal{J} \in \mathcal{P}_k$ belongs to $\mathcal{F}_k$ then every set of $\mathcal{P}_k$ that contains $\mathcal{J}$ belongs to $\mathcal{F}_k$ too.

Generalizing a previous result of Bergsma and Rudas (2002), Bartolucci, Colombi and Forcina (2007) and Cazzaro and Colombi (2006b, 2007) have proved that the *generalized marginal interactions* associated to a *complete* and *hierarchical* family of interactions parameterize the joint distribution of the q categorical variables. Any model obtained by constraining a set of *generalized marginal interactions* associated to a *complete* and *hierarchical* family of interactions is a *HMMM* model that can be fitted with the *hmmm* package.

3 Some Examples and the hmmm package

Many examples of *HMMM* models subject to inequality constraints, together with a discussion of the logits types here introduced can be found in Colombi and Forcina (2001), Bartolucci, Colombi and Forcina (2007) and Cazzaro and Colombi (2006a, 2006b, 2007). Here we give some other examples in order to describe the use of the package *hmmm*. As a first example consider the seemingly unrelated logit regressions represented by the dashed edges graph of figure 5.3(a) of Cox and Wermuth (1996) and here reported in Fig. 1. Under this model the variables A_3 and A_4 are explanatory for the variables A_1 and A_2, A_2 is independent from A_3 given A_4 and A_1 is independent from A_4 given A_3.

The model can be parameterized choosing the complete hierarchical parameterization defined by the marginals $\mathcal{M}_1 = \{3, 4\}$, $\mathcal{M}_2 = \{1, 3, 4\}$,

$\mathcal{M}_3 = \{2, 3, 4\}$, $\mathcal{M}_4 = \{1, 2, 3, 4\}$, and the constraints:

$$\boldsymbol{\eta}_{\{2,3\};\{2,3,4\}}(\boldsymbol{i}_{\{2,3\}}) = 0, \quad \boldsymbol{\eta}_{\{2,3,4\};\{2,3,4\}}(\boldsymbol{i}_{\{2,3,4\}}) = 0,$$
$$\boldsymbol{\eta}_{\{1,4\};\{1,3,4\}}(\boldsymbol{i}_{\{1,4\}}) = 0, \quad \boldsymbol{\eta}_{\{1,3,4\};\{1,3,4\}}(\boldsymbol{i}_{\{1,3,4\}}) = 0.$$
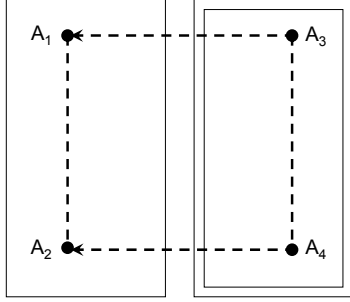

Fig. 1 Figure 5.3(a) of Cox and Wermuth (1996).

If the four categorical variables are ordinal it is sensible to choose logits of type global for A_3 and A_4 within $\mathcal{M}_1$ and for A_1 and A_2 within $\mathcal{M}_2$, $\mathcal{M}_3$ and $\mathcal{M}_4$. As explained in Bartolucci, Colombi and Forcina (2007), who gave a general description of block recursive models of this type, it is convenient to use logits of type local for the explanatory variables A_3 and A_4 within $\mathcal{M}_2$, $\mathcal{M}_3$ and $\mathcal{M}_4$.

If the four categorical variables are ordinal the following inequality constraints:

$$\boldsymbol{\eta}_{\{2,4\};\{2,3,4\}}(\boldsymbol{i}_{\{2,4\}}) \le 0, \quad \boldsymbol{\eta}_{\{1,3\};\{1,3,4\}}(\boldsymbol{i}_{\{1,3\}}) \le 0,$$

state that the distributions of A_2 conditioned by the explanatory variables are stochastically decreasing with the categories of A_4 and that the conditional distributions of A_1 are stochastically decreasing with the categories of A_3. The problem of testing linear inequality constraints on marginal parameters has been discussed by Dardanoni and Forcina (1998), Colombi and Forcina (2001) and by Bartolucci, Colombi and Forcina (2007).

We now illustrate how to specify and estimate the previous model using the package *hmmm*. First of all the marginal sets within which the interactions are allocated must be defined. Every marginal set must be defined by a list with the two components *marg, types*. The first component *marg* describes the marginal set and the second component *types* describes the logit types that are used to build up the interactions *marg*. So in the case of four variables $marg = c(3, 4)$ describes the marginal bivariate distribution

of A_3, A_4 and $types$=c("marg","marg","g","g") says that variables A_1, A_2 are marginalized and that global logits are used for the remaining two variables. Use "g", "l", "c", "rc" for global logits, local logits, continuation and reverse continuation logits respectively. For recursive logits ("r") and base-line logits ("b") see the next sub-section. Finally the *HMMM* model is specified by the statement *hmmm.model(marg,lev)* where *marg* is the list of all the marginal sets previously defined and *lev* contains the number of categories of the variables.

$$library(hmmm)$$
$$marg34 < -list(marg = c(3, 4), types = c("marg", "marg", "g", "g"))$$
$$marg134 < -list(marg = c(1, 3, 4), types = c("g", "marg", "l", "l"))$$
$$marg234 < -list(marg = c(2, 3, 4), types = c("marg", "g", "l", "l"))$$
$$marg1234 < -list(marg = c(1, 2, 3, 4), types = c("g", "g", "l", "l"))$$
$$marglist < -list(marg34, marg134, marg234, marg1234)$$
$$my.model < -hmmm.model(marg = marglist, lev = c(3, 3, 2, 4))$$

A more convenient way to define the list *marglist* is given by the function *marg.list* that is the first six statements can be replaced by:

$$library(hmmm)$$
$$marglist < -c("m - m - g - g", "g - m - l - l", "m - g - l - l", "g - g - l - l")$$
$$marglist < -marg.list(marglist, mflag = "m")$$

The statement $d < -hmmm.model.summary(my.model)$ produces the following output.

inter.	marg.	logit	npar	start	end
3	34	g	1	1	1
4	34	g	3	2	4
34	34	gg	3	5	7
1	134	g	2	8	9
13	134	gl	2	10	11
14	134	gl	6	12	17
134	134	gll	6	18	23
2	234	g	2	24	25
23	234	gl	2	26	27
24	234	gl	6	28	33
234	234	gll	6	34	39
12	1234	gg	4	40	43
123	1234	ggl	4	44	47
124	1234	ggl	12	48	59
1234	1234	ggll	12	60	71

In the first column the interactions are reported, the second column describes the marginal distribution where the interactions are defined. For every interaction type the third column describes the logit type used for every variable in the interaction set and the fourth column reports the number of parameters. The last two columns give the position of the parameters in the vector where all the 71 interactions are arranged. The previous model is a saturated model. Note that within a marginal distribution are defined all the interactions, between the variables in the marginal set, that have not been defined within previous marginal distributions.

The function *marg.list* can be used only in this case that is when within a marginal distribution are defined all the interactions that have not been defined within previous marginal distributions and in particular it cannot be used to define inequalities as in the case of the following statements. This is so because in this case a list of interactions (*int*) must be supplied in the definition of a marginal set.

The equality constraints described at the beginning of the section state that the parameters in the positions: 12,13,14,15,16,17,18,19,20,21,22,23, 26,27,34,35,36,37,38,39, are null.

The following statements define a model with the inequality and equality constraints defined in the first part of this section.

$$marg234ineq < -$$
$$list(marg = c(2,3,4), int = list(c(2,4)), types = c(\text{``marg''}, \text{``g''}, \text{``l''}, \text{``l''}))$$
$$marg134ineq$$
$$< -list(marg = c(1,3,4), int = list(c(1,3)), types = c(\text{``g''}, \text{``marg''}, \text{``l''}, \text{``l''}))$$
$$ineq < -list(marg234ineq, marg134ineq)$$
$$sel < -c(12:23, 26:27, 34:39)$$
$$XX < -diag(1,71)$$
$$XX < -XX[, -sel]$$
$$model.constr < -$$
$$hmmm.model(marg = marglist, lev = c(3,3,2,4),$$
$$dismarg = ineq, X = XX, D = diag(-1,8))$$

Note that in the previous statement the matrix XX defining the equality constraints is assigned to the argument X and the list *ineq* of the interactions subject to inequality constraints is assigned to the argument *dismarg.* The argument D is used to turn the eight inequalities from non-negativity constraints (the default) to non-positivity constraints.

The R function *hmmm.model* creates also the R functions needed to compute the Link function, that is the vector $\boldsymbol{\eta} = \boldsymbol{C}\ln(\boldsymbol{M\pi})$ of the *generalized marginal interactions,* and its derivative with respect to the vector $\boldsymbol{\pi}$ of the joint probabilities. The R functions that compute the equality and inequality constraints and their derivatives with respect to the joint probabilities are also created. These R functions are required to compute the

M.L. estimates of the parameters of the *HMMM* model and to solve many inferential problems.

To show how the previous model can be estimated on real data, we analyze the Madsen data (Madsen, 1976) concerning rental property residents who are classified according to the following categorical variables: *Influence on apartment management* (variable 1), *Satisfaction* (variable 2), *Contact with other residents* (variable 3), and *Housing type* (variable 4). The eight contingency tables concerning the joint distributions of *Influence* and *Satisfaction*, for every level of *Contact* and *Housing*, are given in Table 4.

Table 4 Madsen data.

CONTACT		Low			High		
SATISFACTION		Low	Medium	High	Low	Medium	High
HOUSING	INFLUENCE						
Tower	Low	21	21	28	14	19	37
Blocks	Medium	34	22	36	17	23	40
	High	10	11	36	3	5	23
Apart-	Low	61	23	17	78	46	43
ments	Medium	43	35	40	48	45	86
	High	26	18	54	15	25	62
Atrium	Low	13	9	10	20	23	20
houses	Medium	8	8	12	10	22	24
	High	6	7	9	7	10	21
Terraced	Low	18	6	7	57	23	13
Houses	Medium	15	13	13	31	21	13
	High	7	5	11	5	6	13

Let y be the vector of the sample frequencies arranged in such a way that the modality index i_j of the $j - th$ variable, $j = 1, 2, 3$ runs faster than the indexes to its right. The model can be estimated by the following statements:

$$mod.fit < -hmmm.mlfit(y, model.constr, noineq = FALSE)$$
$$descr fit < -$$
$$hmmm.model.summary(model.constr, mod.fit)$$
$$mph.summary(mod.fit)$$

The statements *mph.summary* and *hmmm.model.summary* give different summaries of the estimated model; *hmmm.model.summary* is specific to *HMMM* models and gives detailed informations on the interactions used, the marginal distribution and the stratum to which they belong together with their point estimates. The Lang's function *mph.summary* is more general and can be used also for *MPH* models.

If $noineq = FALSE$ is omitted or replaced by $noineq = TRUE$ in *hmmm.mlfit* the inequality constraints are ignored and the model is estimated only under the equality constraints:

$$mod.fit.noineq < -hmmm.mlfit(y, model.constr, noineq = TRUE)$$

To estimate the model where inequalities are turned into equalities the following statements must be used:

$$sel < -c(10 : 23, 26 : 39)$$
$$XXX < -diag(1, 71)$$
$$XXX < -XXX[, -sel]$$
$$model.noineq < -$$
$$hmmm.model(marg = marg.list, lev = c(3, 3, 2, 4), dismarg = ineq, X = XXX)$$
$$mod.fit.null < -hmmm.mlfit(y, model.noineq)$$

To obtain the log-likelihood ratio statistics and the simulated *p-values* of the chi-bar square distribution to test *mod.fit.noineq* against *mod.fit* (test type B, Silvapulle and Sen, 2005) and to test *mod.fit* against the model *mod.fit.null* where the inequalities are turned into equalities (test type A, Silvapulle and Sen, 2005), the following statements must be used:

$$p < -chibar(model = model.constr, nullfit = mod.fit.null,$$
$$disfit = mod.fit, satfit = mod.fit.noineq)$$
$$chibar.summary(p)$$

The statement *chibar.summary(p)* gives the output of Table 5.

Table 5 Chi-bar simulated p-values.

	test	p-value
testA	69.754	0
testB	8.06	0.104

As the likelihood ratio statistics to test the model *mod.fit.noineq* against the saturated model is $G^2 = 40.22, (df = 20)$ and as testB is not significative, the model *mod.fit* can be retained. The model previously introduced states that *Influence* is independent from *Housing type* given *Contact* and that *Satisfaction* is independent from *Contact* given *Housing type*. The inequalities state that the distribution of *Influence* given high contact is stochastically smaller than the one given low contact and that the distribution of *Satisfaction* given *Housing type* becomes stochastically smaller when

we move from the top of the table to the bottom. This is quite surprising
because it means that in better houses residents are generally less satisfied
and that residents with low contact generally think to have good influence
more easily than the ones with high contact.

Finally we observe that when the marginal sets are not specified $hmmm.model$
adopts the Glonek and McCullagh (1995) parameterizations with interactions
tions of type local as shown in the following example:

$$library(hmmm)$$
$$model < -hmmm.model(lev = c(3, 3, 2, 4))$$
$$d < -hmmm.model.summary(model)$$

The statement $d < -hmmm.model.summary(model)$ now produces the
following output:

inter.	marg.	logit	npar	start	end
1	1	l	2	1	2
2	2	l	2	3	4
3	3	l	1	5	5
4	4	l	3	6	8
12	12	ll	4	9	12
13	13	ll	2	13	14
23	23	ll	2	15	16
14	14	ll	6	17	22
24	24	ll	6	23	28
34	34	ll	3	29	31
123	123	lll	4	32	35
124	124	lll	12	36	47
134	134	lll	6	48	53
234	234	lll	6	54	59
1234	1234	llll	12	60	71

In the Glonek and McCullagh (1995) parameterizations all the possible
marginal distribution sets are considered and in every marginal distribution
the only interaction set which is defined is the marginal set itself.

When a saturated log-linear model is required the statement

$$lgnmod < -loglin.model(c(3, 3, 2, 4))$$

can be used. The statement $d < -hmmm.model.summary(lgnmod)$ pro-
duces an output with the same interactions given in the previous output
which now are of baseline type and are all defined in the joint distribu-
tions. The function $loglin.model$ can also be used to specify non-saturated
hierarchical log-linear models, if the list of the maximal interaction sets of
non-null interactions is assigned to the argument int. As an alternative the
list of the minimal interaction sets of null interactions can be specified. Fi-
nally local interactions can be used instead of the baseline ones which are
the default.

3.1 Models with recursive and baseline logits

Interactions defined by baseline logits ("b") and recursive logits ("r") are not automatically generated. This is so because recursive logits depend on the *Coherent Complete Hierarchy of Sets* which is used and baseline logits depend on the category chosen as the reference one. Note that in section two we assumed that the reference category was the first one but this is not necessary and even arbitrary for non ordinal variables. Recursive logits and odds ratios have been introduced by Cazzaro and Colombi (2006b, 2007) with special reference to monotone dependence hypotheses and the baseline logits are the only type of logits that can be used with non-ordinal variables. To define these interactions the argument *cocacontr* of the function *hmmm.model* must be used. The argument *cocacontr* must be equal to a list of zero-one matrices; one for every variable of the joint distribution. The first $r_j - 1$ rows of the $j - th$ matrix must describe the indicator functions of the $r_j - 1$ sets $\mathcal{B}_j(m_j, 0)$ and the remaining $r_j - 1$ rows must describe the indicator functions of the sets $\mathcal{B}_j(m_j, 1)$. In the list *cocacontr* matrices, corresponding to variables to which a "b" or "r" logits is never assigned, can be defined arbitrarily (a simple 0 will do) and once a matrix is assigned to a variable A_j it will be used in every marginal distribution where A_j do not have "l", "g", "r", "rc" logits. So the logit type of a variable have not to be the same across different marginal distribution but the reference category or the *CCHS* cannot be changed. This also implies that "b" logits and "r" logits cannot be assigned to the same variable in different marginal distributions. In the following example recursive logits are used for A_1 and baseline logits are used for A_2. Both variables are supposed to have four categories. The non-minimal sets of the coherent complete hierarchy of sets, used for A_1, are the set of all the four categories, the set of the last two categories and the set of the first two categories. The reference category of A_2 is assumed to be the first one.

$$library(hmmm)$$
$$marginal1 < -list(marg = c(1), types = c(\text{``}r\text{''}, \text{``}marg\text{''}))$$
$$marginal2 < -list(marg = c(2), types = c(\text{``}marg\text{''}, \text{``}b\text{''}))$$
$$marginal12 < -list(marg = c(1, 2), types = c(\text{``}r\text{''}, \text{``}b\text{''}))$$
$$marg.list < -list(marginal1, marginal2, marginal12)$$
$$model < -hmmm.model(marg = marg.list, dismarg = dism, lev = c(4, 4),$$
$$cocacontr = list(matrix($$
$$c(1, 1, 0, 0,$$
$$1, 0, 0, 0,$$
$$0, 0, 1, 0,$$
$$0, 0, 1, 1,$$
$$0, 1, 0, 0,$$
$$0, 0, 0, 1),$$

$$6, 4, byrow = TRUE),$$
$$matrix($$
$$c(1, 0, 0, 0,$$
$$1, 0, 0, 0$$
,
$$1, 0, 0, 0,$$
$$0, 1, 0, 0,$$
$$0, 0, 1, 0,$$
$$0, 0, 0, 1),$$
$$6, 4, byrow = TRUE)))$$

A saturated model is defined by the previous statements and the statement $d < -hmmm.model.summary(model)$ now produces the following output.

inter.	marg.	logit	npar	start	end
1	1	r	3	1	3
2	2	b	3	4	6
12	12	rb	9	7	15

Non-saturated models can be defined as shown in the previous example or in the following section.

4 Hypotheses testing on equality and inequality constraints in presence of covariates

When the data comes from s different strata or populations let $\boldsymbol{\eta}_t$, $t = 1, 2, ..., s$, be the vector where the interactions associated to a complete hierarchical family of interactions are arranged. The complete hierarchical family of interactions and the used logit types are supposed to be the same in all s populations. Let $\boldsymbol{\eta}$ be the vector obtained by staking the $\boldsymbol{\eta}_t$'s. The vector $\boldsymbol{\eta}$ may be explicitly written in matrix form as

$$\boldsymbol{\eta} = \boldsymbol{C} \log(\boldsymbol{M}\boldsymbol{\pi}), \tag{2}$$

where the rows of $\boldsymbol{C}$ are contrasts and $\boldsymbol{M}$ is a matrix of zeros and ones which sums the probabilities to obtain the necessary marginal probabilities. The definition of these matrices was given by Colombi and Forcina (2001) and was extended by Cazzaro and Colombi (2006b) in order to include recursive type interactions.

It is interesting to test equality and inequality constraints on the parameter vector $\boldsymbol{\eta}$ under product multinomial sampling. Constraints defined by equalities may be useful for testing conditional independencies or additive effects of covariates in marginal tables or marginal homogeneity hypotheses. Inequality constraints can be useful for verifying monotone dependence or positive association hypotheses. A *HMMM* model is a model of the form:

$$\boldsymbol{\eta} = \boldsymbol{X}\boldsymbol{\beta},$$

where $\boldsymbol{\beta}$ is a vector of unknown parameters and where the elements of the matrix $\boldsymbol{X}$ are known functions of the covariates that describe the strata.

HMMM models are special cases of Lang's (2005) *HLP* models and the way of estimating $\boldsymbol{\beta}$ and testing linear equality is discussed in Lang (2005). The way of testing inequality constraints on $\boldsymbol{\eta}$ is the same as that described by Colombi and Forcina (2001). We only recall that in presence of inequality constraints the likelihood ratio test statistics has an asymptotic chi-bar square distribution. In the next sub-sections the use of the package *hmmm* is described.

4.1 Examples

In the following example *education* (variable 2) with three categories and *jobs* (variable 1) with three categories are the dependent variables. Two explanatory variables *age* (two levels) and *area* (four levels) define eight strata. In the order of the 8 strata the levels of *age* run faster than the ones of *area*. The following statements read the data and define a *HMMM* model as in section 2.3. Note that now the argument *strata* is used in *hmmm.model* to specify the number of strata.

$$library(hmmm)$$
$$y < -c(413, 3184, 2281, 1, 42, 28, 0,$$
$$12, 5, 79, 197, 95, 1, 3, 2,$$
$$0, 0, 0, 52, 330, 274, 2, 9,$$
$$12, 0, 5, 5, 15, 21, 16,$$
$$1, 0, 2, 0, 1, 0, 523, 2945,$$
$$1632, 38, 210, 103, 14, 88, 42,$$
$$113, 302, 81, 7, 11, 2, 2, 3,$$
$$0, 560, 2330, 1365, 49, 139, 81,$$
$$19, 64, 24, 116, 253, 56, 6,$$
$$8, 0, 1, 0, 1)$$
$$y < -matrix(y, 72, 1)$$
$$marg1 < -list(marg = c(2), types = c(\text{``marg''}, \text{``c''}))$$
$$marg12 < -list(marg = c(1, 2), types = c(\text{``g''}, \text{``l''}))$$
$$marg.list < -list(marg1, marg12)$$
$$marg12bis < -list(marg = c(1, 2), int = list(c(1, 2)), types = c(\text{``g''}, \text{``l''}))$$
$$model < -hmmm.model(marg = marg.list, lev = c(3, 3),$$
$$dismarg = list(marg12bis), strata = 8, D = diag(-1, 32))$$

The argument *dismarg* specify that 32 odds ratios (four for every stratum) are constrained by inequalities. If $\boldsymbol{C}_1 \log(\boldsymbol{M}_1 \boldsymbol{\pi})$ is the vector of interactions

defined by *dismarg* and if $D=\boldsymbol{K}$ (the default for the argument D is the identity matrix) then the inequalities constraints are $\boldsymbol{KC}_1\log(\boldsymbol{M}_1\boldsymbol{\pi}) > \boldsymbol{0}$. So in the previous example the 32 o.r. are constrained to be negative. Remember that *hmmm.model* creates the R functions to compute the link function (2), the inequality (equality) constraints and their Hessians.

The parameters of *model* are not subject to equality constraints because by default the matrix $\boldsymbol{X}$ is the identity matrix. The following statement *createX.mat* builds the design matrix $\boldsymbol{X}$ starting from a list of three model formula (one for the marginal logits of variable 2, one for the logits of variable 1 and one for the odds ratios). The function *createX.mat* creates the R functions needed to compute the equality constraints function $\boldsymbol{U}'\boldsymbol{C}\log(\boldsymbol{M}\boldsymbol{\pi})$ and its Hessian. Note that $\boldsymbol{U}$ is a matrix such that $\boldsymbol{U}'\boldsymbol{X} = \boldsymbol{0}$. The first argument of *createX.mat* must have been created by *hmmm.model*.

The following list *al* of model formula defines an additive effect of the covariates on the logits. For the odds ratios interaction between the covariates is allowed but main effects and interaction effects are common to all four odds ratios in a stratum. In the following statement f_0 is a conventional factor name that must be used to define specific or generic effects of the covariates on the interactions. For example "$f\_0 + area + age$" applied to the logits states additive (generic) effects that are common to all logits and "$f\_0 : area + f\_0 : age$" additive effects which are specific to every logit. The formula "$f\_0 + area * age$" applied to the odds ratios states generic and non-additive effects of the covariates.

$$al < -list(" \ 1 + f\_0 * area * age - f\_0 : area : age - area : age",$$
$$" \ 1 + f\_0 * area * age - f\_0 : area : age - area : age"," \ 1 + f\_0 + area * age")$$
$$modelsat < -create.XMAT(model, Formula = al, strata = c(2, 4),$$
$$fnames = c("age", "area"))$$

The following statements define the model where inequalities are turned into equalities. The formula "*zero*" applied to the odds ratios states that all the odds ratios are null.

$$al < -list(" \ 1 + f\_0 * area * age - f\_0 : area : age - area : age",$$
$$" \ 1 + f\_0 * area * age - f\_0 : area : age - area : age", "zero")$$
$$modelnull < -$$
$$create.XMAT(model, Formula = al, strata = c(2, 4),$$
$$fnames = c("age", "area"))$$

The following statements estimate the reference model *modelsat* without inequality constraints, under the hypothesis that all the o.r. are negative

and under the hypothesis of stochastic independence within every stratum (inequalities are turned into equalities). Finally the model with negative association is tested against the model without inequalities and the model of stochastic independence is tested against the model of negative association by *hmmm.chibar*:

$$mods < -hmmm.mlfit(y, modelsat)$$
$$mod0 < -hmmm.mlfit(y, modelnull)$$
$$moddis < -hmmm.mlfit(y, modelsat, noineq = FALSE)$$
$$P < -hmmm.chibar(model = modelsat, nullfit = mod0,$$
$$disfit = moddis, satfit = mods)$$

The statement *ddis< −hmmm.model.summary(modellosat,moddis)* gives the output of Table 6 where the estimates of the *generalized marginal interactions* classified by type and strata are reported.

Table 6 Estimates of the generalized interactions of *moddis* (Likelihood ratio statistics=59.71, df=33).

inter.	marg.	logit	MLE_1	MLE_2	MLE_3	MLE_4	MLE_5	MLE_6	MLE_7	MLE_8
2	2	c	-4.16935	-4.774285	-2.916174	-3.521108	-2.335015	-2.939949	-2.438111	-3.043046
2	2	c	-1.468873	-2.20091	-0.787007	-1.519044	-0.882648	-1.614685	-0.929127	-1.661163
1	12	g	2.557718	1.532542	2.366263	1.341087	2.185841	1.160666	1.89829	0.873114
1	12	g	-0.447417	-1.277674	-0.291748	-1.122005	-0.759915	-1.590172	-0.760161	-1.590419
12	12	gl	-0.02426	-0.02426	-0.02426	-0.02426	-0.062032	-0.517051	-0.207142	-0.550114
12	12	gl	0	0	0	0	-0.037772	-0.492792	-0.182882	-0.525854
12	12	gl	0	0	0	0	-0.037772	-0.492792	-0.182882	-0.525854
12	12	gl	-0.108454	-0.108454	-0.108454	-0.108454	-0.146227	-0.601246	-0.291337	-0.634308

The variances of these estimates together with the estimate of the parameters vector $\boldsymbol{\beta}$ are given by *mph.summary*. The output of *chibar.summary(P)* is reported in Table 7. It is about the tests on the inequality constraints that is of the test of *mod0* against *moddis* (Type A) and on the test of *moddis* against *mods* (Type B).

Table 7 Chi-bar simulated p-values.

	test	p-value
testA	13.106	0.031
testB	3.005	0.773

According to the values of the Likelihood ratio statistics=59.71 (df=33) and of the Type B test, the proposed model for the joint distribution of *education* and *jobs* cannot be rejected.

Going back to the Madsen data of Table 4 let us assume that *Housing type* is a covariate that defines four sub-populations or strata. In this case we

have to model the joint distribution of *Influence* (variable 1), *Satisfaction* (variable 2) and *Contact* (variable 3) within each of the four housing types. We choose to define the logits within the marginal distributions while the other interactions are defined within the joint distributions. Global logits are assigned to *Influence* and *Satisfaction*, while the choice of the logit type is irrelevant for the dicothomous variable *Contact*. Moreover null third order interactions are assumed within every stratum and the second order interactions are assumed to be equal across the strata. Only the marginal logits are allowed to depend on the housing type but the effect of *Housing type* is assumed to be of generic form. We are also interested in testing for positive association between *Influence* and *Satisfaction* conditionally to *Contact* and *Housing type* and between *Contact* and *Satisfaction* conditionally on *Influence* and *Housing type*. The model is defined and estimated with and without inequalities by the following statements:

$$marg < -c(``g - m - m", ``m - g - m", ``m - m - l", ``g - g - l")$$
$$marg < -marg.list(marg, mflag = ``m")$$
$$ineq < -list(marg = c(1, 2, 3), int = list(c(1, 2), c(2, 3)),$$
$$types = c(``g", ``g", ``l"))$$
$$model < -hmmm.model(marg = marg, lev = c(3, 3, 2),$$
$$strata = 4, dismarg = list(ineq), D = diag(1, 24))$$
$$al < -list(`` - 1 + f_0 + Ho", `` - 1 + f_0 + Ho", `` - 1 + Ho",$$
$$`` - 1 + f_0", `` - 1 + f_0", `` - 1 + f_0",$$
$$``zero")$$
$$modelsat < -create.XMAT(model, Formula = al,$$
$$strata = 4, fnames = c("Ho"))$$
$$mod < -hmmm.mlfit(y, modelsat)$$
$$descr < -hmmm.model.summary(modelsat, mod, T)$$
$$modineq < -hmmm.mlfit(y, modelsat, noineq = FALSE)$$

The following statements estimate the previous model with inequalities turned into equalities and perform type A and type B tests on the inequalities:

$$al < -list(`` - 1 + f_0 + Ho", `` - 1 + f_0 + Ho", `` - 1 + Ho",$$
$$``zero", `` - 1 + f_0", ``zero", ``zero")$$
$$modelsat < -create.XMAT(model, Formula = al,$$
$$strata = c(4), fnames = c("Ho"))$$
$$mod0 < -hmmm.mlfit(y, modelsat)$$
$$P < -hmmm.chibar(model = model,$$
$$nullfit = mod0, disfit = modineq, satfit = mod)$$

The output of *hmmm.model.summary(modelsat,mod,T)* is given in Table 8.

Table 8 Estimates of the generalized interactions of *mod* (Likelihood ratio statistics=55.84, df=46).

inter.	marg.	logit	MLE_1	MLE_2	MLE_3	MLE_4
1	1	g	0.572172	0.62986	0.491943	0.184951
1	1	g	-1.138108	-1.08042	-1.218337	-1.525329
2	2	g	1.125874	0.651487	0.829899	0.065262
2	2	g	0.003234	-0.471152	-0.29274	-1.057377
3	3	l	-0.192824	0.350023	0.645332	0.646015
12	123	gg	0.798405	0.798405	0.798405	0.798405
12	123	gg	0.912703	0.912703	0.912703	0.912703
12	123	gg	0.827487	0.827487	0.827487	0.827487
12	123	gg	0.996916	0.996916	0.996916	0.996916
13	123	gl	-0.376978	-0.376978	-0.376978	-0.376978
13	123	gl	-0.517317	-0.517317	-0.517317	-0.517317
23	123	gl	0.395319	0.395319	0.395319	0.395319
23	123	gl	0.24272	0.24272	0.24272	0.24272
123	123	ggl	0	0	0	0
123	123	ggl	0	0	0	0
123	123	ggl	0	0	0	0
123	123	ggl	0	0	0	0

The output given by statement *chibar.summary(P)* is given in Table 9. From Table 9 we see that testB is null because the interactions 12 and 13 are already positive in the model that has been estimated disregarding the inequality constraints. It is also interesting to note that the interactions 13 are negative.

Table 9 Chi-bar simulated p-values.

	test	p-value
testA	1.171145e+02	0
testB	0	1

This confirm the findings of the previous section but here the association between *Influence* and *Contact* is conditioned by *Housing type* and *Satisfaction*, while in the previous section it was conditioned only by *Housing type*. The generic effects (baseline) of *Housing type* on the logits of *Satisfaction*, which can be printed by *mph.summary*, are -0.474386535 for apartments, -0.295974576 for atrium houses and -1.060611663 for terraced houses. This, with the exception of atrium houses, confirms the findings of the previous

section. How to obtain a formal test on this form of monotone dependence is an exercise left to the reader.

5 Estimating MPH and HLP models under inequality constraints

The package *hmmm* can also be used to estimate more general *MPH*, *HLP* models (Lang 2004, 2005) under inequality and equality constraints. To this purpose the function *mphineq* and *chibar* must be directly used and no functions like *hmmm.model*, *hmmm.mlfit*, *hmmm.chibar* to specify, estimate and test these models are available.

Appendix: Coherent Complete Hierarchy of Sets

By *Coherent Complete Hierarchy of Sets* (CCHS) $\mathcal{G}(\mathcal{A})$ a family of non-empty subsets of $\mathcal{A} = \{a_1, a_2, ..., a_r\}$ is intended. It is characterized by the following properties:

- $\mathcal{A} \in \mathcal{G}(\mathcal{A})$;
- $\{a_i\} \in \mathcal{G}(\mathcal{A})$, $i = 1, 2, \ldots, r$;
- if $P, Q \in \mathcal{G}(\mathcal{A})$ then $P \cap Q \in \{P, Q, \emptyset\}$;
- if M_m is a non-minimal set (or *node*), according to the inclusion order, of $\mathcal{G}(\mathcal{A})$ then:

$$M_m = \mathcal{B}(a_{i(m)}, 0) \cup \mathcal{B}(a_{i(m)}, 1),$$
$$\mathcal{B}(a_{i(m)}, 0) \in \mathcal{G}(\mathcal{A}), \quad \mathcal{B}(a_{i(m)}, 1) \in \mathcal{G}(\mathcal{A}),$$
$$i(m) = \sup\{i : a_i \in \mathcal{B}(a_{i(m)}, 0)\},$$
$$i(m) + 1 = \inf\{i : a_i \in \mathcal{B}(a_{i(m)}, 1)\},$$
$$m = 1, 2, \ldots, r - 1.$$

Note that the last property defines a bijective mapping between the sets M_m and the first $r-1$ elements of $\mathcal{A}$. Furthermore, from the same property it follows that each set M_m contains contiguous categories.

A CCHS is completely described by its non minimal sets. The numbering of the non-minimal sets M_m is assumed such that

$$m > n \Rightarrow (M_n \supset M_m \quad \vee \quad \sup\{i : a_i \in M_m\} < \inf\{i : a_i \in M_n\}).$$

References

1. Bartolucci, F., Colombi, R., and Forcina, A., An extended class of marginal link functions for modelling contingency tables by equality and inequality constraints, Statistica Sinica, (2007) to appear.
2. Bergsma, W. P. and Rudas, T., Marginal models for categorical data, Annals of Statistics **30**, (2002) 140-159.

3. Cazzaro, M., Colombi, R., Maximum likelihood inference for log-linear models subject to constraints of double monotone dependence, Statistical Methods and Applications **15**, (2006a) 177-190.

4. Cazzaro, M., Colombi, R., A new family of interactions for modelling contingency tables, (2006b) submitted.

5. Cazzaro, M., Colombi, R., Modelling two way contingency tables with recursive logits and odds ratios, (2007) submitted.

6. Colombi, R., Forcina, A., Marginal regression models for the analysis of positive association of ordinal response variables, Biometrika **88**, (2001) 1007-1019.

7. Cox, D., R., Wermuth, N., *Multivariate dependencies: models, analysis and interpretation* (Chapman & Hall, London 1996).

8. Dardanoni, V., Forcina, A., A unified approach to likelihood inference on stochastic orderings in a nonparametric context, Journal of the American Statistical Association **93**, (1998) 1112-1123.

9. Douglas, R., Fienberg, S., E., Lee, M. T., Sampson, A., R., Whitaker L., R., Poisitve dependence concepts for ordinal contingency tables. In *Topiccs in statistical dependence*, Block, H., W., Sampson, A., R., Sanits, T., H., editors, Institute of mathematical statistics, lecture notes, monograph series; Haywar, California **16**, (1990) 189-202.

10. Glonek, G., F., V., McCullagh, P., Multivariate logistic models for Contingency Tables. Journal of the Royal Statistical Society, B, **57**, (1995) 533-546.

11. Lang, J., B., Multinomial Poisson Homogeneous Models for Contingency Tables. The Annnals of Statistics **32**, (2004) 340-383.

12. Lang, J., B., Homogeneous Linear Predictor Models for Contingency Tables. Journal of the American Statistical Association **100**, (2005) 121-134.

13. Madsen M., Statistical Analysis of Multiple Contingency Tables-Two Examples. Scand. J. Statis. **3**,(1976) 97-106.

14. Silvapulle, M.J., Sen, P.K., *Constrained statistical inference* (Wiley, New Jersey 2005).