Hierarchical Multinomial Marginal models (HMM)

Roberto Colombi

Dipartimento di Ingegneria dell' Informazione e Metodi Matematici
Università di Bergamo
Italy

Riunione Scientifica SIS
Arcavacata June 2008

Outline

Outline

Outline

1. Hierarchical Multinomial Marginal (HMM) models
   - Introduction
   - Bartolucci-Colombi-Forcina (BCF) 2007 parametrization and generalized marginal interactions
   - MPH models HMM models and ML estimation
2. Block recursive models (Chain models)
   - Definitions
   - BCF parametrization: three chain components example
3. Advantages of HMM models

Outline

Outline

Outline

1. Hierarchical Multinomial Marginal (HMM) models
   - Introduction
   - Bartolucci-Colombi-Forcina (BCF) 2007 parametrization and generalized marginal interactions
   - MPH models HMM models and ML estimation
2. Block recursive models (Chain models)
   - Definitions
   - BCF parametrization: three chain components example
3. Advantages of HMM models

Outline

Outline

Outline

Log-linear Models

In the log-linear parametrization all the interactions are **contrasts of logarithms of joint probabilities** and this is the main reason why this parametrization is not convenient to express hypotheses on marginal distributions or to model ordered categorical data.

HMM models

The *Hierarchical Multinomial Marginal* (HMM) models are based on parameters, called *generalized marginal interactions* (Bartolucci-Colombi-Forcina (2007)), which are **contrasts of logarithms of sums of probabilities defined within different marginal distributions**.

Log-linear Models

In the log-linear parametrization all the interactions are **contrasts of logarithms of joint probabilities** and this is the main reason why this parametrization is not convenient to express hypotheses on marginal distributions or to model ordered categorical data.

HMM models

The *Hierarchical Multinomial Marginal* (HMM) models are based on parameters, called *generalized marginal interactions* (Bartolucci-Colombi-Forcina (2007)), which are **contrasts of logarithms of sums of probabilities defined within different marginal distributions**.

Special cases

- The *Log-linear Models* are HMM models where the interactions are defined within the joint probability function.
- The Bergsma-Rudas (2002) *Marginal Models* are HMM models where the interactions are of log-linear type but are defined in different marginal distributions.
- The Glonek-McCullagh (1995) *Multivariate Logit Models* are HMM models where the parameters are the highest order interactions that can be defined within each of the marginal distributions.

Special cases

- The *Log-linear Models* are HMM models where the interactions are defined within the joint probability function.
- The Bergsma-Rudas (2002) *Marginal Models* are HMM models where the interactions are of log-linear type but are defined in different marginal distributions.
- The Glonek-McCullagh (1995) *Multivariate Logit Models* are HMM models where the parameters are the highest order interactions that can be defined within each of the marginal distributions.

Special cases

- The *Log-linear Models* are HMM models where the interactions are defined within the joint probability function.
- The Bergsma-Rudas (2002) *Marginal Models* are HMM models where the interactions are of log-linear type but are defined in different marginal distributions.
- The Glonek-McCullagh (1995) *Multivariate Logit Models* are HMM models where the parameters are the highest order interactions that can be defined within each of the marginal distributions.

Outline

Where and how the Bartolucci-Colombi-Forcina (BCF)interactions are defined

- where: *hierarchical family of marginal sets, complete hierarchical family of interaction sets*
- how: standard log-linear interaction defined on *lumped marginal tables*

Where and how the Bartolucci-Colombi-Forcina (BCF)interactions are defined

- where: *hierarchical family of marginal sets, complete hierarchical family of interaction sets*
- how: standard log-linear interaction defined on *lumped marginal tables*

Marginal sets and Marginal distributions

Marginal sets

A subset, of the set $\{A_j,\ j = 1, 2, ..., q\}$ of all the variables, that defines a marginal distribution is denoted by the set $\mathcal{M}$ of the indices of the corresponding variables. The set $\mathcal{M}$ is called *marginal set*

Marginal distributions

The marginal distribution associated to a marginal set $\mathcal{M}$ is called $\mathcal{M}$-marginal distribution. The set $\mathcal{Q} = \{1, \ldots, q\}$ refers to the joint distribution

examples: $\{3\}$-marginal distribution; $\{1, 3\}$-marginal distribution....

Marginal sets and Marginal distributions

Marginal sets

A subset, of the set $\{A_j,\ j = 1, 2, ..., q\}$ of all the variables, that defines a marginal distribution is denoted by the set $\mathcal{M}$ of the indices of the corresponding variables. The set $\mathcal{M}$ is called *marginal set*

Marginal distributions

The marginal distribution associated to a marginal set $\mathcal{M}$ is called $\mathcal{M}$-marginal distribution. The set $\mathcal{Q} = \{1, \ldots, q\}$ refers to the joint distribution

examples: $\{3\}$-marginal distribution; $\{1, 3\}$-marginal distribution....

Families of marginal sets

hierarchical family of marginal sets

An ordered family $\mathcal{H} = \{\mathcal{M}_1, ..., \mathcal{M}_s\}$ of marginal sets is called *hierarchical family of marginal sets* if it is ordered coherently with the partial order of inclusion. ($\mathcal{M}_k$ is not a subset of $\mathcal{M}_h$ for every $h < k$, $k = 1, 2, ..., s$).

- example $\mathcal{H} = \{\{1\}, \{2\}, \{1, 2\}\}$
- example $\mathcal{H} = \{\{1\}, \{2\}, \{1, 2\}, \{1, 2, 3\}, \{1, 2, 4\}, \{1, 2, 3, 4\}\}$

Families of marginal sets

hierarchical family of marginal sets

An ordered family $\mathcal{H} = \{\mathcal{M}_1, ..., \mathcal{M}_s\}$ of marginal sets is called *hierarchical family of marginal sets* if it is ordered coherently with the partial order of inclusion. ($\mathcal{M}_k$ is not a subset of $\mathcal{M}_h$ for every $h < k$, $k = 1, 2, ..., s$).

- example $\mathcal{H} = \{\{1\}, \{2\}, \{1, 2\}\}$
- example $\mathcal{H} = \{\{1\}, \{2\}, \{1, 2\}, \{1, 2, 3\}, \{1, 2, 4\}, \{1, 2, 3, 4\}\}$

Families of marginal sets

hierarchical family of marginal sets

An ordered family $\mathcal{H} = \{\mathcal{M}_1, ..., \mathcal{M}_s\}$ of marginal sets is called *hierarchical family of marginal sets* if it is ordered coherently with the partial order of inclusion. ($\mathcal{M}_k$ is not a subset of $\mathcal{M}_h$ for every $h < k$, $k = 1, 2, ..., s$).

- example $\mathcal{H} = \{\{1\}, \{2\}, \{1, 2\}\}$
- example $\mathcal{H} = \{\{1\}, \{2\}, \{1, 2\}, \{1, 2, 3\}, \{1, 2, 4\}, \{1, 2, 3, 4\}\}$

Interaction sets

interaction sets

For every $\mathcal{M} \in \mathcal{H}$ the subsets $\mathcal{I} \subseteq \mathcal{M}$ are called interaction sets.

families of interaction sets

- let $\mathcal{P}_k$ be the family of all non-empty subsets of $\mathcal{M}_k$
- let $\mathcal{F}_k \subseteq \mathcal{P}_k$ be the family of *interaction sets* allocated within the $\mathcal{M}_k$-marginal distribution

Interaction sets

interaction sets

For every $\mathcal{M} \in \mathcal{H}$ the subsets $\mathcal{I} \subseteq \mathcal{M}$ are called interaction sets.

families of interaction sets

- let $\mathcal{P}_k$ be the family of all non-empty subsets of $\mathcal{M}_k$
- let $\mathcal{F}_k \subseteq \mathcal{P}_k$ be the family of *interaction sets* allocated within the $\mathcal{M}_k$-marginal distribution

Interaction sets

interaction sets

For every $\mathcal{M} \in \mathcal{H}$ the subsets $\mathcal{I} \subseteq \mathcal{M}$ are called interaction sets.

families of interaction sets

- let $\mathcal{P}_k$ be the family of all non-empty subsets of $\mathcal{M}_k$
- let $\mathcal{F}_k \subseteq \mathcal{P}_k$ be the family of *interaction sets* allocated within the $\mathcal{M}_k$-marginal distribution

Interaction sets

interaction sets

For every $\mathcal{M} \in \mathcal{H}$ the subsets $\mathcal{I} \subseteq \mathcal{M}$ are called interaction sets.

families of interaction sets

- let $\mathcal{P}_k$ be the family of all non-empty subsets of $\mathcal{M}_k$
- let $\mathcal{F}_k \subseteq \mathcal{P}_k$ be the family of *interaction sets* allocated within the $\mathcal{M}_k$-marginal distribution

Complete hierarchical family of interaction sets

Given a hierarchical family $\mathcal{H} = \{\mathcal{M}_1, ..., \mathcal{M}_s\}$ of marginal sets the ordered list of interaction families:

$$\mathcal{F}_1, ..., \mathcal{F}_s$$

is called *complete hierarchical family of interaction sets* if it satisfies the following (completeness, non redundancy and hierarchy) conditions:

- $\bigcup_{h=1}^{s} \mathcal{F}_h = \{\mathcal{I} : \mathcal{I} \neq \emptyset, \mathcal{I} \subseteq \mathcal{Q}\}$
- $\mathcal{F}_1 = \mathcal{P}_1$
- $\mathcal{F}_k = \mathcal{P}_k \setminus \bigcup_{h<k} \mathcal{F}_h$.

Complete hierarchical family of interaction sets

Given a hierarchical family $\mathcal{H} = \{\mathcal{M}_1, ..., \mathcal{M}_s\}$ of marginal sets the ordered list of interaction families:

$$\mathcal{F}_1, ..., \mathcal{F}_s$$

is called *complete hierarchical family of interaction sets* if it satisfies the following (completeness, non redundancy and hierarchy) conditions:

- $\bigcup_{h=1}^{s} \mathcal{F}_h = \{\mathcal{I} : \mathcal{I} \neq \emptyset, \mathcal{I} \subseteq \mathcal{Q}\}$
- $\mathcal{F}_1 = \mathcal{P}_1$
- $\mathcal{F}_k = \mathcal{P}_k \setminus \bigcup_{h<k} \mathcal{F}_h$.

Complete hierarchical family of interaction sets

Given a hierarchical family $\mathcal{H} = \{\mathcal{M}_1, ..., \mathcal{M}_s\}$ of marginal sets the ordered list of interaction families:

$$\mathcal{F}_1, ..., \mathcal{F}_s$$

is called *complete hierarchical family of interaction sets* if it satisfies the following (completeness, non redundancy and hierarchy) conditions:

- $\bigcup_{h=1}^{s} \mathcal{F}_h = \{\mathcal{I} \,:\, \mathcal{I} \neq \emptyset,\, \mathcal{I} \subseteq \mathcal{Q}\}$
- $\mathcal{F}_1 = \mathcal{P}_1$
- $\mathcal{F}_k = \mathcal{P}_k \setminus \bigcup_{h<k} \mathcal{F}_h$.

Complete hierarchical family of interaction sets

Given a hierarchical family $\mathcal{H} = \{\mathcal{M}_1, ..., \mathcal{M}_s\}$ of marginal sets the ordered list of interaction families:

$$\mathcal{F}_1, ..., \mathcal{F}_s$$

is called *complete hierarchical family of interaction sets* if it satisfies the following (completeness, non redundancy and hierarchy) conditions:

- $\bigcup_{h=1}^{s} \mathcal{F}_h = \{\mathcal{I} : \mathcal{I} \neq \emptyset, \mathcal{I} \subseteq \mathcal{Q}\}$
- $\mathcal{F}_1 = \mathcal{P}_1$
- $\mathcal{F}_k = \mathcal{P}_k \setminus \bigcup_{h<k} \mathcal{F}_h$.

three categorical variables (Glonek McCullagh interaction sets)

interaction sets	marginal sets
A_1	A_1
A_2	A_2
A_1, A_2	A_1, A_2
A_3	A_3
A_1, A_3	A_1, A_3
A_2, A_3	A_2, A_3
A_1, A_2, A_3	A_1, A_2, A_3

Note. Marginal sets and interaction sets are defined as sets of integers because this simplifies the notation for the interaction parameters. However in the examples sets of variables are more convenient.

Generalized marginal interactions

Any *generalized marginal interaction* is defined by

- the *interaction set* $\mathcal{I}$ of the variables interacting with one another,
- the $\mathcal{M}$*-marginal distribution* where it is defined, $\mathcal{M} \supseteq \mathcal{I}$,
- the logit types assigned to the variables of the $\mathcal{M}$-distribution.

Generalized marginal interactions

Any *generalized marginal interaction* is defined by

- the *interaction set* $\mathcal{I}$ of the variables interacting with one another,
- the $\mathcal{M}$*-marginal distribution* where it is defined, $\mathcal{M} \supseteq \mathcal{I}$,
- the logit types assigned to the variables of the $\mathcal{M}$-distribution.

Generalized marginal interactions

Any *generalized marginal interaction* is defined by

- the *interaction set* $\mathcal{I}$ of the variables interacting with one another,
- the $\mathcal{M}$*-marginal distribution* where it is defined, $\mathcal{M} \supseteq \mathcal{I}$,
- the logit types assigned to the variables of the $\mathcal{M}$-distribution.

Logits defined on marginal distributions

Given a categorical variable A_1 with r_1 categories and given $r_1 - 1$ pairs of disjoint subsets of categories:

$$\mathcal{B}_1(m_1, 0), \mathcal{B}_1(m_1, 1),\ m_1 = 1, 2, ..., r_1 - 1,$$

the logits, defined on the univariate $\{1\}-$marginal distribution, are the log-probability odds:

Logits

$$\ln \frac{P(A_1 \in \mathcal{B}_1(m_1, 1))}{P(A_1 \in \mathcal{B}_1(m_1, 0))},\ m_1 = 1, 2, ..., r_1 - 1$$

Logits defined on marginal distributions

Given a categorical variable A_1 with r_1 categories and given $r_1 - 1$ pairs of disjoint subsets of categories:

$$\mathcal{B}_1(m_1, 0), \mathcal{B}_1(m_1, 1),\ m_1 = 1, 2, ..., r_1 - 1,$$

the logits, defined on the univariate $\{1\}-$marginal distribution, are the log-probability odds:

Logits

$$\ln \frac{P(A_1 \in \mathcal{B}_1(m_1, 1))}{P(A_1 \in \mathcal{B}_1(m_1, 0))},\ m_1 = 1, 2, ..., r_1 - 1$$

Logits defined on marginal distributions

Given a categorical variable A_1 with r_1 categories and given $r_1 - 1$ pairs of disjoint subsets of categories:

$$\mathcal{B}_1(m_1, 0), \mathcal{B}_1(m_1, 1), \; m_1 = 1, 2, ..., r_1 - 1,$$

the logits, defined on the univariate $\{1\}-$marginal distribution, are the log-probability odds:

Logits

$$\ln \frac{P(A_1 \in \mathcal{B}_1(m_1, 1))}{P(A_1 \in \mathcal{B}_1(m_1, 0))}, \; m_1 = 1, 2, ..., r_1 - 1$$

baseline, local, global, continuation aggregation criteria

aggregation criterion	$\mathcal{B}_1(m_1, 0)$	$\mathcal{B}_1(m_1, 1)$
base-line	$\{a_{1,1}\}$	$\{a_{1,m_1+1}\}$
local	$\{a_{1,m_1}\}$	$\{a_{1,m_1+1}\}$
global	$\{a_{1,1}, \ldots, a_{1,m_1}\}$	$\{a_{1,m_1+1}, \ldots, a_{1,r_1}\}$
continuation	$\{a_{1,m_1}\}$	$\{a_{1,m_1+1}, \ldots, a_{1,r_1}\}$

Reference category set ($m_1 = 1$): $\mathcal{B}_1(1, 0) = \{a_{1,1}\}$

log odds ratios defined on bivariate distributions

Given a logit type for the categorical variable A_1 and a logit type for the categorical variable A_2 we can compute the probabilities:

$$p_{\{1,2\}}(h_1, h_2; m_1, m_2) = pr(A_1 \in \mathcal{B}_1(m_1, h_1), A_2 \in \mathcal{B}_2(m_2, h_2)),$$
$$m_1 = 1, 2, ..., r_1 - 1, m_2 = 1, 2, ..., r_2 - 1, h_1 = 0, 1, h_2 = 0, 1.$$

a family of generalized log-odds ratios, defined on the $\{1,2\}-$marginal distribution of A_1, A_2, is composed by the standard log-odds ratios computed in the tables called *lumped table*:

	$\mathcal{B}_2(m_2, 0)$	$\mathcal{B}_2(m_2, 1)$
$\mathcal{B}_1(m_1, 0)$	$p_{\{1,2\}}(0, 0; m_1, m_2)$	$p_{\{1,2\}}(0, 1; m_1, m_2)$
$\mathcal{B}_1(m_1, 1)$	$p_{\{1,2\}}(1, 0; m_1, m_2)$	$p_{\{1,2\}}(1, 1; m_1, m_2)$

log odds ratios defined on bivariate distributions

Given a logit type for the categorical variable A_1 and a logit type for the categorical variable A_2 we can compute the probabilities:

$$p_{\{1,2\}}(h_1, h_2; m_1, m_2) = pr(A_1 \in \mathcal{B}_1(m_1, h_1), A_2 \in \mathcal{B}_2(m_2, h_2)),$$
$$m_1 = 1, 2, ..., r_1 - 1, m_2 = 1, 2, ..., r_2 - 1, h_1 = 0, 1, h_2 = 0, 1.$$

a family of generalized log-odds ratios, defined on the $\{1, 2\}-$marginal distribution of A_1, A_2, is composed by the standard log-odds ratios computed in the tables called *lumped table*:

	$\mathcal{B}_2(m_2, 0)$	$\mathcal{B}_2(m_2, 1)$
$\mathcal{B}_1(m_1, 0)$	$p_{\{1,2\}}(0, 0; m_1, m_2)$	$p_{\{1,2\}}(0, 1; m_1, m_2)$
$\mathcal{B}_1(m_1, 1)$	$p_{\{1,2\}}(1, 0; m_1, m_2)$	$p_{\{1,2\}}(1, 1; m_1, m_2)$

generalized log odds ratios

- 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.).

generalized log odds ratios

- 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.).

where the interactions are defined?

Responses from a clinical trial comparing three treatments

A_1 : *Outcome* / A_2 : *Treatment*	*D*	*MaD*	*MiD*
P	59	71	80
LD	48	65	77
MD	87	121	194

where are defined the interactions?

Lumped tables

A_1 : *Out.* / A_2 : *Treat.*	*D*	*MaD* ∨ *MiD*	*MaD*	*MiD*	A_2 *Treat.*
P	59	151	71	80	210
LD ∨ *MD*	135	457	186	271	592
LD	48	142	65	77	190
MD	87	315	121	194	402
A_1 *Out.*	194	608	278	351	

generalized marginal interactions

The parameters called *generalized marginal interactions* are such that:

- the previous logits and log odds ratios are special cases
- any interaction is a contrast of the previous logits and log odds ratios computed in different marginal conditional distributions

generalized marginal interactions

The parameters called *generalized marginal interactions* are such that:

- the previous logits and log odds ratios are special cases
- any interaction is a contrast of the previous logits and log odds ratios computed in different marginal conditional distributions

generalized marginal interactions

The parameters called *generalized marginal interactions* are such that:

- the previous logits and log odds ratios are special cases
- any interaction is a contrast of the previous logits and log odds ratios computed in different marginal conditional distributions

Lumped marginal tables

In order to introduce in full generality the *generalized marginal interactions* we introduce the *Lumped marginal tables*

- for every $\mathcal{M} \in \mathcal{H}$ the vector $\underline{\mathrm{m}}_{\mathcal{M}}$ of integers

$$m_j,\ 1 \leq m_j < r_j,\ j \in \mathcal{M}$$

  identifies a $2^{\sharp(M)}$ lumped marginal table $T_{\underline{\mathrm{m}}_{\mathcal{M}}}$
- the probabilities of the marginalized and aggregated table (*Lumped marginal tables*) $T_{\underline{\mathrm{m}}_{\mathcal{M}}}$ are:

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

- $\underline{\mathrm{h}}_{\mathcal{M}}$ is a row vector whose elements h_j, $j \in \mathcal{M}$, are equal to zero or to one
- $\underline{\mathrm{h}}_{\mathcal{M}}$ identifies the different cells of the $2^{\sharp(\mathcal{M})}$ table $T_{\underline{\mathrm{m}}_{\mathcal{M}}}$

Lumped marginal tables

In order to introduce in full generality the *generalized marginal interactions* we introduce the *Lumped marginal tables*

- for every $\mathcal{M} \in \mathcal{H}$ the vector $\underline{m}_{\mathcal{M}}$ of integers

$$m_j,\ 1 \leq m_j < r_j,\ j \in \mathcal{M}$$

  identifies a $2^{\sharp(M)}$ lumped marginal table $T_{\underline{m}_{\mathcal{M}}}$
- the probabilities of the marginalized and aggregated table (*Lumped marginal tables*) $T_{\underline{m}_{\mathcal{M}}}$ are:

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

- $\underline{h}_{\mathcal{M}}$ is a row vector whose elements h_j, $j \in \mathcal{M}$, are equal to zero or to one
- $\underline{h}_{\mathcal{M}}$ identifies the different cells of the $2^{\sharp(\mathcal{M})}$ table $T_{\underline{m}_{\mathcal{M}}}$

Lumped marginal tables

In order to introduce in full generality the *generalized marginal interactions* we introduce the *Lumped marginal tables*

- for every $\mathcal{M} \in \mathcal{H}$ the vector $\underline{m}_{\mathcal{M}}$ of integers

$$m_j,\ 1 \leq m_j < r_j,\ j \in \mathcal{M}$$

identifies a $2^{\sharp(M)}$ lumped marginal table $T_{\underline{m}_{\mathcal{M}}}$

- the probabilities of the marginalized and aggregated table (*Lumped marginal tables*) $T_{\underline{m}_{\mathcal{M}}}$ are:

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

- $\underline{h}_{\mathcal{M}}$ is a row vector whose elements h_j, $j \in \mathcal{M}$, are equal to zero or to one
- $\underline{h}_{\mathcal{M}}$ identifies the different cells of the $2^{\sharp(\mathcal{M})}$ table $T_{\underline{m}_{\mathcal{M}}}$

Lumped marginal tables

In order to introduce in full generality the *generalized marginal interactions* we introduce the *Lumped marginal tables*

- for every $\mathcal{M} \in \mathcal{H}$ the vector $\underline{\mathrm{m}}_{\mathcal{M}}$ of integers

$$m_j,\ 1 \leq m_j < r_j,\ j \in \mathcal{M}$$

  identifies a $2^{\sharp(M)}$ lumped marginal table $T_{\underline{\mathrm{m}}_{\mathcal{M}}}$
- the probabilities of the marginalized and aggregated table (*Lumped marginal tables*) $T_{\underline{\mathrm{m}}_{\mathcal{M}}}$ are:

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

- $\underline{\mathrm{h}}_{\mathcal{M}}$ is a row vector whose elements h_j, $j \in \mathcal{M}$, are equal to zero or to one
- $\underline{\mathrm{h}}_{\mathcal{M}}$ identifies the different cells of the $2^{\sharp(\mathcal{M})}$ table $T_{\underline{\mathrm{m}}_{\mathcal{M}}}$

Lumped marginal tables

In order to introduce in full generality the *generalized marginal interactions* we introduce the *Lumped marginal tables*

- for every $\mathcal{M} \in \mathcal{H}$ the vector $\underline{\mathrm{m}}_{\mathcal{M}}$ of integers

$$m_j, \; 1 \leq m_j < r_j, \; j \in \mathcal{M}$$

  identifies a $2^{\sharp(M)}$ lumped marginal table $T_{\underline{\mathrm{m}}_{\mathcal{M}}}$
- the probabilities of the marginalized and aggregated table (*Lumped marginal tables*) $T_{\underline{\mathrm{m}}_{\mathcal{M}}}$ are:

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

- $\underline{\mathrm{h}}_{\mathcal{M}}$ is a row vector whose elements h_j, $j \in \mathcal{M}$, are equal to zero or to one
- $\underline{\mathrm{h}}_{\mathcal{M}}$ identifies the different cells of the $2^{\sharp(\mathcal{M})}$ table $T_{\underline{\mathrm{m}}_{\mathcal{M}}}$

Lumped marginal tables

In order to introduce in full generality the *generalized marginal interactions* we introduce the *Lumped marginal tables*

- for every $\mathcal{M} \in \mathcal{H}$ the vector $\underline{m}_{\mathcal{M}}$ of integers

$$m_j, \ 1 \leq m_j < r_j, \ j \in \mathcal{M}$$

  identifies a $2^{\sharp(M)}$ lumped marginal table $T_{\underline{m}_{\mathcal{M}}}$
- the probabilities of the marginalized and aggregated table (*Lumped marginal tables*) $T_{\underline{m}_{\mathcal{M}}}$ are:

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

- $\underline{h}_{\mathcal{M}}$ is a row vector whose elements h_j, $j \in \mathcal{M}$, are equal to zero or to one
- $\underline{h}_{\mathcal{M}}$ identifies the different cells of the $2^{\sharp(\mathcal{M})}$ table $T_{\underline{m}_{\mathcal{M}}}$

Generalized marginal interactions

for every marginal set $\mathcal{M}_k \in \mathcal{H}$

for every interaction set $\mathcal{I} \in \mathcal{F}_k$

the interaction multi-index $\underset{\sim}{m}_{\mathcal{I}}$ is a vector of integers with elements m_j, $m_j = 1, 2, ...r_j - 1$, $\forall j \in \mathcal{I}$

every interaction multi-index $\underset{\sim}{m}_{\mathcal{I}}$, is expanded to the vector $\underset{\sim}{m}_{\mathcal{M}_k}$ such that:

$$\underset{\sim}{m}_{\mathcal{M}_k \cap \mathcal{I}} = \underset{\sim}{m}_{\mathcal{I}}, \ \underset{\sim}{m}_{\mathcal{M}_k \setminus \mathcal{I}} = \underset{\sim}{1}_{\mathcal{M}_k \setminus \mathcal{I}}$$

$\underset{\sim}{m}_{\mathcal{M}_k}$ identifies the lumped marginal table $T_{\underset{\sim}{m}_{\mathcal{M}_k}}$ where all the variables belonging to $\mathcal{M}_k$ but not to $\mathcal{I}$ can be fixed at the reference category ($\mathcal{B}_j(1,0) = \{a_{j,1}\} \ \forall j \notin \mathcal{I}$)

the *generalized marginal interactions (BCF-interactions)* $\eta_{\mathcal{I};\mathcal{M}_k}(\underset{\sim}{m}_{\mathcal{I}})$ is defined as a standard baseline $\mathcal{I}$-log-linear interaction in the table $T_{\underset{\sim}{m}_{\mathcal{M}_k}}$

Generalized marginal interactions

for every marginal set $\mathcal{M}_k \in \mathcal{H}$

for every interaction set $\mathcal{I} \in \mathcal{F}_k$

the interaction multi-index $\underline{m}_{\mathcal{I}}$ is a vector of integers with elements m_j, $m_j = 1, 2, \ldots r_j - 1$, $\forall j \in \mathcal{I}$

every interaction multi-index $\underline{m}_{\mathcal{I}}$, is expanded to the vector $\underline{m}_{\mathcal{M}_k}$ such that:

$$\underline{m}_{\mathcal{M}_k \cap \mathcal{I}} = \underline{m}_{\mathcal{I}}, \ \underline{m}_{\mathcal{M}_k \setminus \mathcal{I}} = \underline{1}_{\mathcal{M}_k \setminus \mathcal{I}}$$

$\underline{m}_{\mathcal{M}_k}$ identifies the lumped marginal table $T_{\underline{m}_{\mathcal{M}_k}}$ where all the variables belonging to $\mathcal{M}_k$ but not to $\mathcal{I}$ can be fixed at the reference category ($\mathcal{B}_j(1,0) = \{a_{j,1}\}\ \forall j \notin \mathcal{I}$)

the *generalized marginal interactions (BCF-interactions)* $\eta_{\mathcal{I};\mathcal{M}_k}(\underline{m}_{\mathcal{I}})$ is defined as a standard baseline $\mathcal{I}$-log-linear interaction in the table $T_{\underline{m}_{\mathcal{M}_k}}$

Generalized marginal interactions

for every marginal set $\mathcal{M}_k \in \mathcal{H}$

for every interaction set $\mathcal{I} \in \mathcal{F}_k$

the interaction multi-index $\underline{m}_{\mathcal{I}}$ is a vector of integers with elements $m_j,\ m_j = 1, 2, ...r_j - 1,\ \forall j \in \mathcal{I}$

every interaction multi-index $\underline{m}_{\mathcal{I}}$, is expanded to the vector $\underline{m}_{\mathcal{M}_k}$ such that:

$$\underline{m}_{\mathcal{M}_k \cap \mathcal{I}} = \underline{m}_{\mathcal{I}},\ \underline{m}_{\mathcal{M}_k \setminus \mathcal{I}} = \underline{1}_{\mathcal{M}_k \setminus \mathcal{I}}$$

$\underline{m}_{\mathcal{M}_k}$ identifies the lumped marginal table $T_{\underline{m}_{\mathcal{M}_k}}$ where all the variables belonging to $\mathcal{M}_k$ but not to $\mathcal{I}$ can be fixed at the reference category ($\mathcal{B}_j(1, 0) = \{a_{j,1}\}\ \forall j \notin \mathcal{I}$)

the *generalized marginal interactions (BCF-interactions)* $\eta_{\mathcal{I};\mathcal{M}_k}(\underline{m}_{\mathcal{I}})$ is defined as a standard baseline $\mathcal{I}$-log-linear interaction in the table $T_{\underline{m}_{\mathcal{M}_k}}$

Generalized marginal interactions

for every marginal set $\mathcal{M}_k \in \mathcal{H}$

for every interaction set $\mathcal{I} \in \mathcal{F}_k$

the interaction multi-index $\underline{m}_{\mathcal{I}}$ is a vector of integers with elements $m_j,\ m_j = 1, 2, ...r_j - 1,\ \forall j \in \mathcal{I}$

every interaction multi-index $\underline{m}_{\mathcal{I}}$, is expanded to the vector $\underline{m}_{\mathcal{M}_k}$ such that:

$$\underline{m}_{\mathcal{M}_k \cap \mathcal{I}} = \underline{m}_{\mathcal{I}},\ \underline{m}_{\mathcal{M}_k \setminus \mathcal{I}} = \underline{1}_{\mathcal{M}_k \setminus \mathcal{I}}$$

$\underline{m}_{\mathcal{M}_k}$ identifies the lumped marginal table $T_{\underline{m}_{\mathcal{M}_k}}$ where all the variables belonging to $\mathcal{M}_k$ but not to $\mathcal{I}$ can be fixed at the reference category ($\mathcal{B}_j(1, 0) = \{a_{j,1}\}\ \forall j \notin \mathcal{I}$)

the *generalized marginal interactions (BCF-interactions)* $\eta_{\mathcal{I};\mathcal{M}_k}(\underline{m}_{\mathcal{I}})$ is defined as a standard baseline $\mathcal{I}$-log-linear interaction in the table $T_{\underline{m}_{\mathcal{M}_k}}$

Generalized marginal interactions

for every marginal set $\mathcal{M}_k \in \mathcal{H}$

for every interaction set $\mathcal{I} \in \mathcal{F}_k$

the interaction multi-index $\underline{\mathrm{m}}_{\mathcal{I}}$ is a vector of integers with elements $m_j,\ m_j = 1, 2, \ldots r_j - 1,\ \forall j \in \mathcal{I}$

every interaction multi-index $\underline{\mathrm{m}}_{\mathcal{I}}$, is expanded to the vector $\underline{\mathrm{m}}_{\mathcal{M}_k}$ such that:

$$\underline{\mathrm{m}}_{\mathcal{M}_k \cap \mathcal{I}} = \underline{\mathrm{m}}_{\mathcal{I}},\ \underline{\mathrm{m}}_{\mathcal{M}_k \setminus \mathcal{I}} = \underline{1}_{\mathcal{M}_k \setminus \mathcal{I}}$$

$\underline{\mathrm{m}}_{\mathcal{M}_k}$ identifies the lumped marginal table $T_{\underline{\mathrm{m}}_{\mathcal{M}_k}}$ where all the variables belonging to $\mathcal{M}_k$ but not to $\mathcal{I}$ can be fixed at the reference category ($\mathcal{B}_j(1,0) = \{a_{j,1}\}\ \forall j \notin \mathcal{I}$)

the *generalized marginal interactions (BCF-interactions)* $\eta_{\mathcal{I};\mathcal{M}_k}(\underline{\mathrm{m}}_{\mathcal{I}})$ is defined as a standard baseline $\mathcal{I}$-log-linear interaction in the table $T_{\underline{\mathrm{m}}_{\mathcal{M}_k}}$

Generalized marginal interactions

- for every marginal set $\mathcal{M}_k \in \mathcal{H}$
- for every interaction set $\mathcal{I} \in \mathcal{F}_k$
- the interaction multi-index $\underline{m}_{\mathcal{I}}$ is a vector of integers with elements $m_j,\ m_j = 1, 2, ...r_j - 1,\ \forall j \in \mathcal{I}$
- every interaction multi-index $\underline{m}_{\mathcal{I}}$, is expanded to the vector $\underline{m}_{\mathcal{M}_k}$ such that:

$$\underline{m}_{\mathcal{M}_k \cap \mathcal{I}} = \underline{m}_{\mathcal{I}},\ \underline{m}_{\mathcal{M}_k \setminus \mathcal{I}} = \underline{1}_{\mathcal{M}_k \setminus \mathcal{I}}$$

- $\underline{m}_{\mathcal{M}_k}$ identifies the lumped marginal table $T_{\underline{m}_{\mathcal{M}_k}}$ where all the variables belonging to $\mathcal{M}_k$ but not to $\mathcal{I}$ can be fixed at the reference category ($\mathcal{B}_j(1,0) = \{a_{j,1}\}\ \forall j \notin \mathcal{I}$)
- the *generalized marginal interactions (BCF-interactions)* $\eta_{\mathcal{I};\mathcal{M}_k}(\underline{m}_{\mathcal{I}})$ is defined as a standard baseline $\mathcal{I}$-log-linear interaction in the table $T_{\underline{m}_{\mathcal{M}_k}}$

BCF parametrization and HMM models

- The generalized marginal interactions

$$\eta_{\mathcal{I};\mathcal{M}_k}(\underline{m}_{\mathcal{I}}),\ \forall \mathcal{I} \in \mathcal{F}_k,\ \forall \mathcal{M}_k \in \mathcal{H},\ \forall \underline{m}_{\mathcal{I}}$$

(remember that $\underline{m}_{\mathcal{I}}$ has elements $m_j = 1, 2, ...r_j - 1,\ \forall j \in \mathcal{I}$) associated to a *complete hierarchical family of interaction sets*, parameterize the joint distribution of the q categorical variables. (Bartolucci, Colombi and Forcina (2007)(BCF))

- BCF showed that the vector η of the previous interactions can be written as

$$\eta = \mathbf{C}\ln(\mathbf{M}\pi)$$

- Any model defined by linear constraints on the *generalized marginal interactions* associated to a *complete hierarchical family of interactions* is a HMM model.

BCF parametrization and HMM models

- The generalized marginal interactions

$$\eta_{\mathcal{I};\mathcal{M}_k}(\underline{m}_{\mathcal{I}}),\ \forall \mathcal{I} \in \mathcal{F}_k,\ \forall \mathcal{M}_k \in \mathcal{H},\ \forall \underline{m}_{\mathcal{I}}$$

  (remember that $\underline{m}_{\mathcal{I}}$ has elements $m_j = 1, 2, ...r_j - 1,\ \forall j \in \mathcal{I}$) associated to a *complete hierarchical family of interaction sets*, parameterize the joint distribution of the q categorical variables. (Bartolucci, Colombi and Forcina (2007)(BCF))
- BCF showed that the vector η of the previous interactions can be written as

$$\eta = \mathbf{C} \ln(\mathbf{M}\pi)$$

- Any model defined by linear constraints on the *generalized marginal interactions* associated to a *complete hierarchical family of interactions* is a HMM model.

BCF parametrization and HMM models

- The generalized marginal interactions

$$\eta_{\mathcal{I};\mathcal{M}_k}(\underline{m}_{\mathcal{I}}),\ \forall\mathcal{I}\in\mathcal{F}_k,\ \forall\mathcal{M}_k\in\mathcal{H},\ \forall\underline{m}_{\mathcal{I}}$$

  (remember that $\underline{m}_{\mathcal{I}}$ has elements $m_j = 1, 2, ...r_j - 1,\ \forall j \in \mathcal{I}$) associated to a *complete hierarchical family of interaction sets*, parameterize the joint distribution of the q categorical variables. (Bartolucci, Colombi and Forcina (2007)(BCF))
- BCF showed that the vector η of the previous interactions can be written as

$$\eta = \mathbf{C}\ln(\mathbf{M}\pi)$$

- Any model defined by linear constraints on the *generalized marginal interactions* associated to a *complete hierarchical family of interactions* is a HMM model.

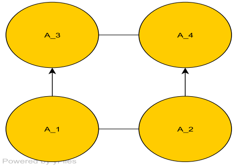
example

$\mathcal{H} = \{\{1\}, \{2\}, \{1,2\}, \{1,2,3\}, \{1,2,4\}, \{1,2,3,4\}\}$

inter.	*marg.*	*type*	*inter.*	*marg.*	*type*
A_1	A_1	g	A_4	A_1, A_2, A_4	g
A_2	A_2	g	A_1, A_4 (=)	A_1, A_2, A_4	lg
A_1, A_2	A_1, A_2	gg	A_2, A_4	A_1, A_2, A_4	lg
			A_1, A_2, A_4 (=)	A_1, A_2, A_4	llg
A_3	A_1, A_2, A_3	g	A_3, A_4 (=)	A_1, A_2, A_3, A_4	ll
A_1, A_3	A_1, A_2, A_3	lg	A_1, A_3, A_4 (=)	A_1, A_2, A_3, A_4	lll
A_2, A_3 (=)	A_1, A_2, A_3	lg	A_2, A_3, A_4 (=)	A_1, A_2, A_3, A_4	lll
A_1, A_2, A_3 (=)	A_1, A_2, A_3	llg	A_1, A_2, A_3, A_4(=)	A_1, A_2, A_3, A_4	llll

$A_3 \perp\!\!\!\perp A_2 | A_1$, $A_4 \perp\!\!\!\perp A_1 | A_2$, $A_3 \perp\!\!\!\perp A_4 | A_1, A_2$

two chain components



Outline

ML estimation in HMM models

- All the linear constraints on generalized marginal interactions can be expressed in the form $\boldsymbol{U} \ln(\boldsymbol{M}\boldsymbol{\pi}) = \mathbf{0}$
- Under the linear constraints, the multinomial log-likelihood can be maximized as shown by Lang (2004,2005) that also gave asymptotic results on ML estimators that are pertinent in this case too; in fact HMM models can be seen as special cases of Lang's MPH models (Colombi Cazzaro 2008).
- the hmmm R-package developed by Colombi Cazzaro (www.unibg.it/pers/?colombi) allows for the presence of linear inequality constraints and has facilities for a friendly specification of the *hierarchical family of marginal sets*, the *complete hierarchical family of interaction sets* and the associated *generalized marginal interactions*.

ML estimation in HMM models

- All the linear constraints on generalized marginal interactions can be expressed in the form $\boldsymbol{U}\ln(\boldsymbol{M}\boldsymbol{\pi}) = \mathbf{0}$
- Under the linear constraints, the multinomial log-likelihood can be maximized as shown by Lang (2004,2005) that also gave asymptotic results on ML estimators that are pertinent in this case too; in fact HMM models can be seen as special cases of Lang's MPH models (Colombi Cazzaro 2008).
- the hmmm R-package developed by Colombi Cazzaro (www.unibg.it/pers/?colombi) allows for the presence of linear inequality constraints and has facilities for a friendly specification of the *hierarchical family of marginal sets*, the *complete hierarchical family of interaction sets* and the associated *generalized marginal interactions*.

ML estimation in HMM models

- All the linear constraints on generalized marginal interactions can be expressed in the form $\boldsymbol{U}\ln(\boldsymbol{M}\boldsymbol{\pi}) = \mathbf{0}$
- Under the linear constraints, the multinomial log-likelihood can be maximized as shown by Lang (2004,2005) that also gave asymptotic results on ML estimators that are pertinent in this case too; in fact HMM models can be seen as special cases of Lang's MPH models (Colombi Cazzaro 2008).
- the hmmm R-package developed by Colombi Cazzaro (www.unibg.it/pers/?colombi) allows for the presence of linear inequality constraints and has facilities for a friendly specification of the *hierarchical family of marginal sets*, the *complete hierarchical family of interaction sets* and the associated *generalized marginal interactions*.

Outline

Block recursive models

motivation

Examples of block recursive models have been examined in Bartolucci (2007 sec. 2.4) and in Colombi and Forcina (2001) because they represent an interesting setting to show how *hierarchical family of marginal sets*, *complete hierarchical family of interaction sets* and the associated *generalized marginal interactions* can be defined in practice.

Block recursive models

A block recursive model is defined by a set of conditional independencies that are encoded by a mixed adicyclic graph or chain graph where vertices represent variables and edges lack of independence

LWF-AMP-CW

Different rules of reading the independencies from the graph lead to different classes of block recursive models (Lauritzen-Wermuth-Frydenberg class, Andersson-Madigan-Perlman class, Cox-Wermuth class)

Block recursive models

A block recursive model is defined by a set of conditional independencies that are encoded by a mixed adicyclic graph or chain graph where vertices represent variables and edges lack of independence

LWF-AMP-CW

Different rules of reading the independencies from the graph lead to different classes of block recursive models (Lauritzen-Wermuth-Frydenberg class, Andersson-Madigan-Perlman class, Cox-Wermuth class)

Block recursive models

A block recursive model is defined by a set of conditional independencies that are encoded by a mixed adicyclic graph or chain graph where vertices represent variables and edges lack of independence

LWF-AMP-CW

Different rules of reading the independencies from the graph lead to different classes of block recursive models (Lauritzen-Wermuth-Frydenberg class, Andersson-Madigan-Perlman class, Cox-Wermuth class)

Block recursive models

A block recursive model is defined by a set of conditional independencies that are encoded by a mixed adicyclic graph or chain graph where vertices represent variables and edges lack of independence

LWF-AMP-CW

Different rules of reading the independencies from the graph lead to different classes of block recursive models (Lauritzen-Wermuth-Frydenberg class, Andersson-Madigan-Perlman class, Cox-Wermuth class)

Chain Graphs

In a mixed adicyclic graph edges can be partitioned into an ordered family of blocks or chain components such that:

- all undirected edges join vertices within the same block
- all directed edges are between vertices in different blocks
- directed edges are oriented in accordance with the ordering of the blocks

Chain Graphs

In a mixed adicyclic graph edges can be partitioned into an ordered family of blocks or chain components such that:

- all undirected edges join vertices within the same block
- all directed edges are between vertices in different blocks
- directed edges are oriented in accordance with the ordering of the blocks

Chain Graphs

In a mixed adicyclic graph edges can be partitioned into an ordered family of blocks or chain components such that:

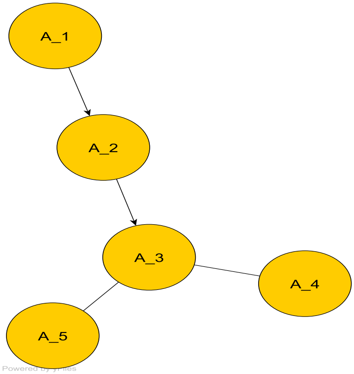
- all undirected edges join vertices within the same block
- all directed edges are between vertices in different blocks
- directed edges are oriented in accordance with the ordering of the blocks

Chain Graphs

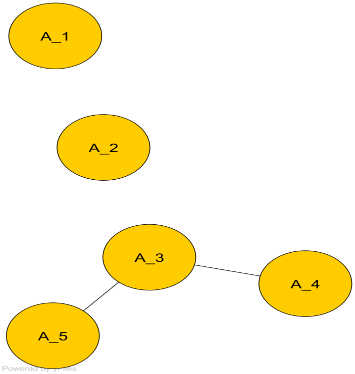
In a mixed adicyclic graph edges can be partitioned into an ordered family of blocks or chain components such that:

- all undirected edges join vertices within the same block
- all directed edges are between vertices in different blocks
- directed edges are oriented in accordance with the ordering of the blocks

three chain components example



chain components



Three types of conditional independencies

- T1: independencies between blocks or chain components
- T2: independencies of sets of variables belonging to a same chain from variables belonging to parent chains
- T3: independencies between variables in a same block

Three types of conditional independencies

- T1: independencies between blocks or chain components
- T2: independencies of sets of variables belonging to a same chain from variables belonging to parent chains
- T3: independencies between variables in a same block

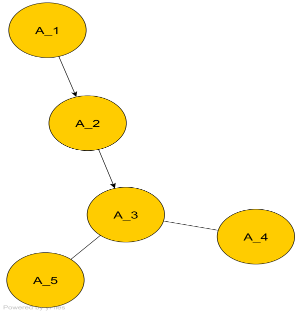
Three types of conditional independencies

- T1: independencies between blocks or chain components
- T2: independencies of sets of variables belonging to a same chain from variables belonging to parent chains
- T3: independencies between variables in a same block

Three types of conditional independencies

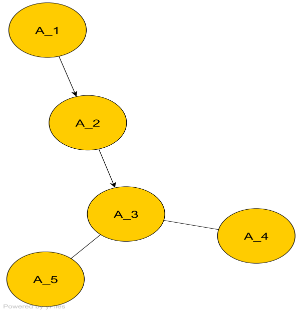
- T1: independencies between blocks or chain components
- T2: independencies of sets of variables belonging to a same chain from variables belonging to parent chains
- T3: independencies between variables in a same block

three chain components example



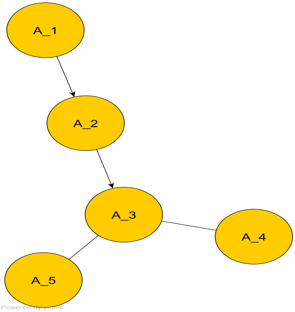
- T1: $A_5 A_4 A_3 \perp\!\!\!\perp A_1 | A_2$
- T2: $A_5 A_4 \perp\!\!\!\perp A_2$
- T3: $A_5 \perp\!\!\!\perp A_4 | A_2$

three chain components example



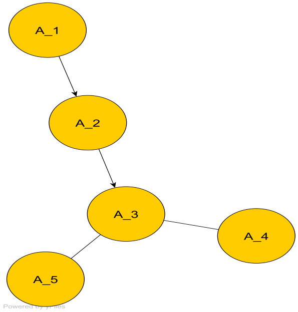
- T1: $A_5A_4A_3 \perp\!\!\!\perp A_1 | A_2$
- T2: $A_5A_4 \perp\!\!\!\perp A_2$
- T3: $A_5 \perp\!\!\!\perp A_4 | A_2$

three chain components example



- T1: $A_5 A_4 A_3 \perp\!\!\!\perp A_1 | A_2$
- T2: $A_5 A_4 \perp\!\!\!\perp A_2$
- T3: $A_5 \perp\!\!\!\perp A_4 | A_2$

three chain components example



- T1: $A_5 A_4 A_3 \perp\!\!\!\perp A_1 | A_2$
- T2: $A_5 A_4 \perp\!\!\!\perp A_2$
- T3: $A_5 \perp\!\!\!\perp A_4 | A_2$

Outline

BCF parametrization: three chain components example

marg.	non null inter.	null inter.	independence
A_1	A_1		
A_1A_2	$A_2,\ A_1A_2$		
A_2A_3	$A_3,\ A_2A_3$		
A_2A_4	A_4	$A_2A_4{\color{red}\bullet}$	
A_2A_5	A_5	$A_2A_5{\color{red}\bullet}$	
$A_2A_3A_4$	$A_3A_4, A_2A_3A_4$		
$A_2A_3A_5$	$A_3A_5, A_2A_3A_5$		
$A_2A_4A_5$		$A_2A_4A_5{\color{blue}\bullet}{\color{red}\bullet}$	${\color{red}A_5A_4 \perp\!\!\!\perp A_2}$
		$A_4A_5{\color{blue}\bullet}$	${\color{blue}A_4 \perp\!\!\!\perp A_5 \mid A_2}$
$A_2A_3A_4A_5$	$A_3A_4A_5,\ A_2A_3A_4A_5$		

BCF parametrization: three chain components example

marg.	non null inter.	null inter.	independence
$A_1A_2A_3A_4A_5$		A_1A_3, A_1A_4, A_1A_5	
		$A_1A_3A_4$, $A_1A_3A_5$	
		$A_1A_4A_5$, $A_1A_3A_4A_5$	
		$A_1A_2A_3$, $A_1A_2A_4$, $A_1A_2A_5$	
		$A_1A_2A_3A_4$, $A_1A_2A_3A_5$	
		$A_1A_2A_4A_5$, $A_1A_2A_3A_4A_5$	$A_5A_4A_3 \perp\!\!\!\perp A_1 \mid A_2$

Advantages of HMM models

- Log-linear models, Glonek-McCullagh models, Bergsma-Rudas models are special cases of HMM models
- HMM models are special cases of Lang's MPH models and the ML-estimation theory and practice can be developed in this general framework
- HMM models are an unified enviroment where well known different types of logits and log- odds ratios can be used as interaction parameters

Advantages of HMM models

- Log-linear models, Glonek-McCullagh models, Bergsma-Rudas models are special cases of HMM models
- HMM models are special cases of Lang's MPH models and the ML-estimation theory and practice can be developed in this general framework
- HMM models are an unified enviroment where well known different types of logits and log- odds ratios can be used as interaction parameters

Advantages of HMM models

- Log-linear models, Glonek-McCullagh models, Bergsma-Rudas models are special cases of HMM models
- HMM models are special cases of Lang's MPH models and the ML-estimation theory and practice can be developed in this general framework
- HMM models are an unified enviroment where well known different types of logits and log- odds ratios can be used as interaction parameters

Advantages of HMM models

- hypotheses on marginal distributions are easily specified by linear constraints
- the presence of ordinal categorical variables can be taken into proper account
- additive effects on response variables can be modelled on different logit scales
- hypotheses of uniform or constant association can be formulated by using different types of log odds ratios
- positive association or monotone dependence hypotheses on different logit scales are given by linear inequality constraints
- HMM models can be used to parametrize block recursive models

Advantages of HMM models

- hypotheses on marginal distributions are easily specified by linear constraints
- the presence of ordinal categorical variables can be taken into proper account
- additive effects on response variables can be modelled on different logit scales
- hypotheses of uniform or constant association can be formulated by using different types of log odds ratios
- positive association or monotone dependence hypotheses on different logit scales are given by linear inequality constraints
- HMM models can be used to parametrize block recursive models

Advantages of HMM models

- hypotheses on marginal distributions are easily specified by linear constraints
- the presence of ordinal categorical variables can be taken into proper account
- additive effects on response variables can be modelled on different logit scales
- hypotheses of uniform or constant association can be formulated by using different types of log odds ratios
- positive association or monotone dependence hypotheses on different logit scales are given by linear inequality constraints
- HMM models can be used to parametrize block recursive models

Advantages of HMM models

- hypotheses on marginal distributions are easily specified by linear constraints
- the presence of ordinal categorical variables can be taken into proper account
- additive effects on response variables can be modelled on different logit scales
- hypotheses of uniform or constant association can be formulated by using different types of log odds ratios
- positive association or monotone dependence hypotheses on different logit scales are given by linear inequality constraints
- HMM models can be used to parametrize block recursive models

Advantages of HMM models

- hypotheses on marginal distributions are easily specified by linear constraints
- the presence of ordinal categorical variables can be taken into proper account
- additive effects on response variables can be modelled on different logit scales
- hypotheses of uniform or constant association can be formulated by using different types of log odds ratios
- positive association or monotone dependence hypotheses on different logit scales are given by linear inequality constraints
- HMM models can be used to parametrize block recursive models

Advantages of HMM models

- hypotheses on marginal distributions are easily specified by linear constraints
- the presence of ordinal categorical variables can be taken into proper account
- additive effects on response variables can be modelled on different logit scales
- hypotheses of uniform or constant association can be formulated by using different types of log odds ratios
- positive association or monotone dependence hypotheses on different logit scales are given by linear inequality constraints
- HMM models can be used to parametrize block recursive models

Applications and Theory

Applications of HMM models to real data

- Bartolucci Forcina (2007)
- Bartolucci Dardanoni Forcina (2001)
- Cazzaro Colombi (2006a, 2006b, 2008)
- Colombi Forcina (2001)

Theoretical backgraund on HMM models

- Bergsma Rudas (2002)
- Bartolucci Colombi Forcina (2007)
- Cazzaro Colombi (2008)
- Colombi Forcina (2001)

Software: HMM models can be easily specified and estimated by the hmmm R-package developed by Cazzaro Colombi (2007) (download it from www.unibg.it/pers/?colombi)

Applications and Theory

Applications of HMM models to real data

- Bartolucci Forcina (2007)
- Bartolucci Dardanoni Forcina (2001)
- Cazzaro Colombi (2006a, 2006b, 2008)
- Colombi Forcina (2001)

Theoretical backgraund on HMM models

- Bergsma Rudas (2002)
- Bartolucci Colombi Forcina (2007)
- Cazzaro Colombi (2008)
- Colombi Forcina (2001)

Software: HMM models can be easily specified and estimated by the hmmm R-package developed by Cazzaro Colombi (2007) (download it from www.unibg.it/pers/?colombi)

Applications and Theory

Applications of HMM models to real data

- Bartolucci Forcina (2007)
- Bartolucci Dardanoni Forcina (2001)
- Cazzaro Colombi (2006a, 2006b, 2008)
- Colombi Forcina (2001)

Theoretical backgraund on HMM models

- Bergsma Rudas (2002)
- Bartolucci Colombi Forcina (2007)
- Cazzaro Colombi (2008)
- Colombi Forcina (2001)

Software: HMM models can be easily specified and estimated by the hmmm R-package developed by Cazzaro Colombi (2007) (download it from www.unibg.it/pers/?colombi)

Applications and Theory

Applications of HMM models to real data

- Bartolucci Forcina (2007)
- Bartolucci Dardanoni Forcina (2001)
- Cazzaro Colombi (2006a, 2006b, 2008)
- Colombi Forcina (2001)

Theoretical backgraund on HMM models

- Bergsma Rudas (2002)
- Bartolucci Colombi Forcina (2007)
- Cazzaro Colombi (2008)
- Colombi Forcina (2001)

Software: HMM models can be easily specified and estimated by the hmmm R-package developed by Cazzaro Colombi (2007) (download it from www.unibg.it/pers/?colombi)

Applications and Theory

Applications of HMM models to real data

- Bartolucci Forcina (2007)
- Bartolucci Dardanoni Forcina (2001)
- Cazzaro Colombi (2006a, 2006b, 2008)
- Colombi Forcina (2001)

Theoretical backgraund on HMM models

- Bergsma Rudas (2002)
- Bartolucci Colombi Forcina (2007)
- Cazzaro Colombi (2008)
- Colombi Forcina (2001)

Software: HMM models can be easily specified and estimated by the hmmm R-package developed by Cazzaro Colombi (2007) (download it from www.unibg.it/pers/?colombi)

Applications and Theory

Applications of HMM models to real data

- Bartolucci Forcina (2007)
- Bartolucci Dardanoni Forcina (2001)
- Cazzaro Colombi (2006a, 2006b, 2008)
- Colombi Forcina (2001)

Theoretical backgraund on HMM models

- Bergsma Rudas (2002)
- Bartolucci Colombi Forcina (2007)
- Cazzaro Colombi (2008)
- Colombi Forcina (2001)

Software: HMM models can be easily specified and estimated by the hmmm R-package developed by Cazzaro Colombi (2007) (download it from www.unibg.it/pers/?colombi)

Applications and Theory

Applications of HMM models to real data

- Bartolucci Forcina (2007)
- Bartolucci Dardanoni Forcina (2001)
- Cazzaro Colombi (2006a, 2006b, 2008)
- Colombi Forcina (2001)

Theoretical backgraund on HMM models

- Bergsma Rudas (2002)
- Bartolucci Colombi Forcina (2007)
- Cazzaro Colombi (2008)
- Colombi Forcina (2001)

Software: HMM models can be easily specified and estimated by the hmmm R-package developed by Cazzaro Colombi (2007) (download it from www.unibg.it/pers/?colombi)

Applications and Theory

Applications of HMM models to real data

- Bartolucci Forcina (2007)
- Bartolucci Dardanoni Forcina (2001)
- Cazzaro Colombi (2006a, 2006b, 2008)
- Colombi Forcina (2001)

Theoretical backgraund on HMM models

- Bergsma Rudas (2002)
- Bartolucci Colombi Forcina (2007)
- Cazzaro Colombi (2008)
- Colombi Forcina (2001)

Software: HMM models can be easily specified and estimated by the hmmm R-package developed by Cazzaro Colombi (2007) (download it from www.unibg.it/pers/?colombi)

Applications and Theory

Applications of HMM models to real data

- Bartolucci Forcina (2007)
- Bartolucci Dardanoni Forcina (2001)
- Cazzaro Colombi (2006a, 2006b, 2008)
- Colombi Forcina (2001)

Theoretical backgraund on HMM models

- Bergsma Rudas (2002)
- Bartolucci Colombi Forcina (2007)
- Cazzaro Colombi (2008)
- Colombi Forcina (2001)

Software: HMM models can be easily specified and estimated by the hmmm R-package developed by Cazzaro Colombi (2007) (download it from www.unibg.it/pers/?colombi)

Applications and Theory

Applications of HMM models to real data

- Bartolucci Forcina (2007)
- Bartolucci Dardanoni Forcina (2001)
- Cazzaro Colombi (2006a, 2006b, 2008)
- Colombi Forcina (2001)

Theoretical backgraund on HMM models

- Bergsma Rudas (2002)
- Bartolucci Colombi Forcina (2007)
- Cazzaro Colombi (2008)
- Colombi Forcina (2001)

Software: HMM models can be easily specified and estimated by the hmmm R-package developed by Cazzaro Colombi (2007) (download it from www.unibg.it/pers/?colombi)

Applications and Theory

Applications of HMM models to real data

- Bartolucci Forcina (2007)
- Bartolucci Dardanoni Forcina (2001)
- Cazzaro Colombi (2006a, 2006b, 2008)
- Colombi Forcina (2001)

Theoretical backgraund on HMM models

- Bergsma Rudas (2002)
- Bartolucci Colombi Forcina (2007)
- Cazzaro Colombi (2008)
- Colombi Forcina (2001)

Software: HMM models can be easily specified and estimated by the hmmm R-package developed by Cazzaro Colombi (2007) (download it from www.unibg.it/pers/?colombi)