Hierarchical models for random permutations

Villő Csiszár

Department of Probability and Statistics
Eötvös Loránd University, Budapest

XXVIII International Seminar on Stability Problems for Stochastic Models

Outline

1. The Models
   - Motivation of the Models
   - Conditional Independence
   - Markov Bases
2. Results
   - L-decomposable model
   - bi-L-decomposable model
   - Quasi-independence model

Outline

Hierarchical Models for Categorical Data.

- $I = I_1 \times \cdots \times I_n$ is a finite sample space.
- $X = (X(1), \ldots, X(n))$ is a random vector taking values in I.
- The distribution of X is $P(X = x) = p(x)$.
- If $x \in I$ and $A \subseteq [n]$, then $x(A) = (x(i) : i \in A)$.
- $\mathcal{A}$ is a collection of subsets of $[n]$.

Definition

The hierarchical model with generator-set $\mathcal{A}$ is the collection of distributions p of form

$$p(x) = \prod_{A \in \mathcal{A}} \theta_A(x(A)) \quad \forall x \in I.$$

Random Permutations.

- S_n is the set of all permutations of $[n] = \{1, \ldots, n\}$.
- $\Pi : \Omega \to S_n$ is a random permutation with $p(\pi) = P(\Pi = \pi)$.
- $\pi = \pi(1), \ldots, \pi(n)$.

Example

Word association data. 129 college students ranked the words *1-benediction, 2-instrument, 3-score, 4-solo, 5-suit* according to the strength of association with the target word *song*.
24 orderings observed, the two most frequent being 42315 (34 times) and 42135 (24 times).

Hierarchical Models for Permutations.

1st idea: Let $I = [n]^n$, apply hierarchical models for categorical data. Disadvantage: many structural zeros!

Our idea: Generalize hierarchical models!

- Generators: product partitions $\mathcal{R} \times \mathcal{C}$ of $[n] \times [n]$, i.e. $\mathcal{R} = (R_1, \ldots, R_r)$ and $\mathcal{C} = (C_1, \ldots, C_c)$.
- Marginals: $r \times c$ matrix $\pi(\mathcal{R} \times \mathcal{C})$, whose ijth element is

$$(\pi(\mathcal{R} \times \mathcal{C}))_{ij} = |\{1 \leq k \leq n : k \in R_i, \pi(k) \in C_j\}|.$$

For $A = \{a_1, \ldots, a_j\} \subseteq [n]$, the marginal $\pi(A)$ is equivalent to $\pi(\mathcal{R} \times \mathcal{C})$ with $\mathcal{R} = (\{a_1\}, \ldots, \{a_j\}, A^c)$ and $\mathcal{C} = (\{1\}, \ldots, \{n\})$, so traditional hierarchical models fit into this new frame as well.

Hierarchical Models for Permutations.

1st idea: Let $I = [n]^n$, apply hierarchical models for categorical data. Disadvantage: many structural zeros!

Our idea: Generalize hierarchical models!

- Generators: product partitions $\mathcal{R} \times \mathcal{C}$ of $[n] \times [n]$, i.e. $\mathcal{R} = (R_1, \ldots, R_r)$ and $\mathcal{C} = (C_1, \ldots, C_c)$.
- Marginals: $r \times c$ matrix $\pi(\mathcal{R} \times \mathcal{C})$, whose ijth element is

$$(\pi(\mathcal{R} \times \mathcal{C}))_{ij} = |\{1 \leq k \leq n : k \in R_i, \pi(k) \in C_j\}|.$$

For $A = \{a_1, \ldots, a_j\} \subseteq [n]$, the marginal $\pi(A)$ is equivalent to $\pi(\mathcal{R} \times \mathcal{C})$ with $\mathcal{R} = (\{a_1\}, \ldots, \{a_j\}, A^c)$ and $\mathcal{C} = (\{1\}, \ldots, \{n\})$, so traditional hierarchical models fit into this new frame as well.

Hierarchical Models for Permutations.

1st idea: Let $I = [n]^n$, apply hierarchical models for categorical data. Disadvantage: many structural zeros!
Our idea: Generalize hierarchical models!

- Generators: product partitions $\mathcal{R} \times \mathcal{C}$ of $[n] \times [n]$, i.e. $\mathcal{R} = (R_1, \ldots, R_r)$ and $\mathcal{C} = (C_1, \ldots, C_c)$.
- Marginals: $r \times c$ matrix $\pi(\mathcal{R} \times \mathcal{C})$, whose ijth element is

$$(\pi(\mathcal{R} \times \mathcal{C}))_{ij} = |\{1 \leq k \leq n : k \in R_i, \pi(k) \in C_j\}|.$$

For $A = \{a_1, \ldots, a_j\} \subseteq [n]$, the marginal $\pi(A)$ is equivalent to $\pi(\mathcal{R} \times \mathcal{C})$ with $\mathcal{R} = (\{a_1\}, \ldots, \{a_j\}, A^c)$ and $\mathcal{C} = (\{1\}, \ldots, \{n\})$, so traditional hierarchical models fit into this new frame as well.

Hierarchical Models for Permutations.

1st idea: Let $I = [n]^n$, apply hierarchical models for categorical data. Disadvantage: many structural zeros!
Our idea: Generalize hierarchical models!

- Generators: product partitions $\mathcal{R} \times \mathcal{C}$ of $[n] \times [n]$, i.e. $\mathcal{R} = (R_1, \ldots, R_r)$ and $\mathcal{C} = (C_1, \ldots, C_c)$.
- Marginals: $r \times c$ matrix $\pi(\mathcal{R} \times \mathcal{C})$, whose ijth element is

$$(\pi(\mathcal{R} \times \mathcal{C}))_{ij} = |\{1 \leq k \leq n : k \in R_i, \pi(k) \in C_j\}|.$$

For $A = \{a_1, \ldots, a_j\} \subseteq [n]$, the marginal $\pi(A)$ is equivalent to $\pi(\mathcal{R} \times \mathcal{C})$ with $\mathcal{R} = (\{a_1\}, \ldots, \{a_j\}, A^c)$ and $\mathcal{C} = (\{1\}, \ldots, \{n\})$, so traditional hierarchical models fit into this new frame as well.

Hierarchical Models for Permutations.

1st idea: Let $I = [n]^n$, apply hierarchical models for categorical data. Disadvantage: many structural zeros!
Our idea: Generalize hierarchical models!

- Generators: product partitions $\mathcal{R} \times \mathcal{C}$ of $[n] \times [n]$, i.e. $\mathcal{R} = (R_1, \ldots, R_r)$ and $\mathcal{C} = (C_1, \ldots, C_c)$.
- Marginals: $r \times c$ matrix $\pi(\mathcal{R} \times \mathcal{C})$, whose ijth element is

$$(\pi(\mathcal{R} \times \mathcal{C}))_{ij} = |\{1 \leq k \leq n : k \in R_i, \pi(k) \in C_j\}|.$$

For $A = \{a_1, \ldots, a_j\} \subseteq [n]$, the marginal $\pi(A)$ is equivalent to $\pi(\mathcal{R} \times \mathcal{C})$ with $\mathcal{R} = (\{a_1\}, \ldots, \{a_j\}, A^c)$ and $\mathcal{C} = (\{1\}, \ldots, \{n\})$, so traditional hierarchical models fit into this new frame as well.

Graphical Models for Categorical Data.

L-decomposable Model

Definition (Luce (1959), Critchlow et al. (1991))

A random ordering Π is L-decomposable, if

$$p(\pi) = \prod_{k=1}^{n} p(\pi(k) \mid \{\pi(k), \ldots, \pi(n)\}).$$

Equivalently, for every k, the first k and last $n-k$ elements of Π are conditionally independent, given the set of the first k elements.

L-decomposable distributions form a hierarchical model with generators $\mathcal{R}_i \times \mathcal{C}$, $i = 2, \ldots, n-1$, where

$$\mathcal{R}_i = (\{1, \ldots, i-1\}, \{i\}, \{i+1, \ldots, n\}) \text{ and } \mathcal{C} = (\{1\}, \ldots, \{n\}).$$

Bi-L-decomposable Model

Rankings can also be L-decomposable, but this depends on the labelling of the alternatives. If there is a natural labelling, then the model makes sense.
Or: If Π is a pairing of two labelled sets, then Π and Π^{-1} are symmetric.

Definition

The random permutation Π is called bi-L-decomposable, if both Π and Π^{-1} are L-decomposable.

The bi-L-decomposable model is also characterized by a set of conditional independence statements.
Question: Is it a hierarchical model?

Quasi-independence Model

Definition

The distribution p is quasi-independent, if it can be written as $p(\pi) = \prod_{k=1}^{n} \theta_k(\pi(k))$.

Quasi-independent distributions form a hierarchical model with generators $\mathcal{R}_i \times \mathcal{R}_j$, $i, j = 1, \ldots, n$, where $\mathcal{R}_i = (\{i\}, [n] \setminus \{i\})$.

Question: Can they be characterized by conditional independence statements?

Markov Bases

Make Titles Informative.

Make Titles Informative.

Make Titles Informative.

Summary

- The first main message of your talk in one or two lines.
- The second main message of your talk in one or two lines.
- Perhaps a third message, but not more than that.

For Further Reading I

A. Author.
Handbook of Everything.
Some Press, 1990.

S. Someone.
On this and that.
Journal of This and That, 2(1):50–100, 2000.