Markov tree models and L-cumulants

Piotr Zwiernik
Mittag-Leffler Institute,
Stockholm

ESI, Vienna,
29 March 2011

Main themes

- Tree cumulants and tree models.
- L-cumulants.

Algebraic statistical model

- $X \in \{1, \ldots, m\}$, $p = (p_1, \ldots, p_m)$ distribution
- p identified with a point in

$$\Delta_{m-1} = \{p \in \mathbb{R}^m : \sum_i p_i = 1,\, p_i \geq 0\}.$$

- a statistical model is any family of points in Δ_{m-1}
- algebraic statistical model $p : \Theta \to \Delta$ and p is a polynomial map
  - e.g. $X \perp\!\!\!\perp Y$ then $p_{ij} = s_i t_j$ and hence $p : \Delta_{m-1} \times \Delta_{n-1} \to \Delta_{mn-1}$

The multivariate cumulants

- let $X = (X_1, \ldots, X_n)$ such that $\mu_{\alpha_1 \cdots \alpha_n} = \mathbb{E}(X_1^{\alpha_1} \cdots X_n^{\alpha_n})$.
- $M_X(t) = \sum_{\alpha \geq 0} \frac{\mu_\alpha}{\alpha!} t^\alpha$ is the m.g.f, $t^\alpha = t_1^{\alpha_1} \cdots t_n^{\alpha_n}$, $\alpha! = \alpha_1! \cdots \alpha_n!$
- $K_X(t) = \log M_X(t) = \sum_{\alpha \geq 0} \frac{k_\alpha}{\alpha!} t^\alpha$ is the c.g.f

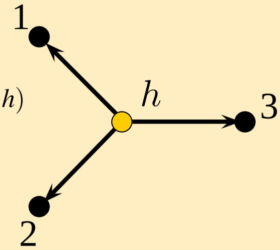
The setting

- Here $X = (X_1, \ldots, X_n) \in \{0,1\}^n$.
- $P = [p_\alpha]$ s.t. $\alpha \in \{0,1\}^n$

- Square-free moments: $m_\beta = \sum_{\alpha \geq \beta} p_\alpha$ for $\beta \in \{0,1\}^n$
- change notation: $m_\beta \mapsto m_B$ for $B \subseteq [n] = \{1, \ldots, n\}$.
- e.g. $n = 3$ then $\mu_{110} \mapsto \mu_{12}$, $\mu_{111} \mapsto \mu_{123}$

The tripod tree model

- $X_1 \perp\!\!\!\perp X_2 \perp\!\!\!\perp X_3 | H$, $X = (X_1, X_2, X_3) \sim p$
- $\mathbb{P}(X = x) = \sum_h \mathbb{P}(H) \prod_{i=1}^{3} \mathbb{P}(X_i = x_i | H = h)$
- $$p_{ijk}(t, a^{(0)}, a^{(1)}, b^{(0)}, b^{(1)}, c^{(0)}, c^{(1)}) =$$ $$= (1-t) a_i^{(0)} b_j^{(0)} c_k^{(0)} + t a_i^{(1)} b_j^{(1)} c_k^{(1)}.$$
- e.g. $a_i^{(0)} = \mathbb{P}(X_1 = i | H = 0)$
- $a_0^{(i)} + a_1^{(i)} = 1$, $b_0^{(i)} + b_1^{(i)} = 1$, $c_0^{(i)} + c_1^{(i)} = 1$ for $i = 0, 1$.



Change of coordinates

- cumulants: $k_1, k_2, k_3, k_{12}, k_{13}, k_{23}, k_{123}$.
- $$[p_{ijk} : i,j,k = 0,1] \xleftrightarrow{1-1} [k_I : I \subseteq [3]].$$

$$k_1 = a_1^{(0)} + t(a_1^{(1)} - a_1^{(0)}), \ldots$$

$$k_{12} = t(1-t)(a_1^{(1)} - a_1^{(0)})(b_1^{(1)} - b_1^{(0)}), \ldots$$

$$k_{123} = t(1-t)(1-2t)(a_1^{(1)} - a_1^{(0)})(b_1^{(1)} - b_1^{(0)})(c_1^{(1)} - c_1^{(0)}).$$

- define $\bar{a} = a_1^{(1)} - a_1^{(0)}$, $\bar{b} = b_1^{(1)} - b_1^{(0)}$, $\bar{c} = c_1^{(1)} - c_1^{(0)}$ and $s = 1 - 2t$ then

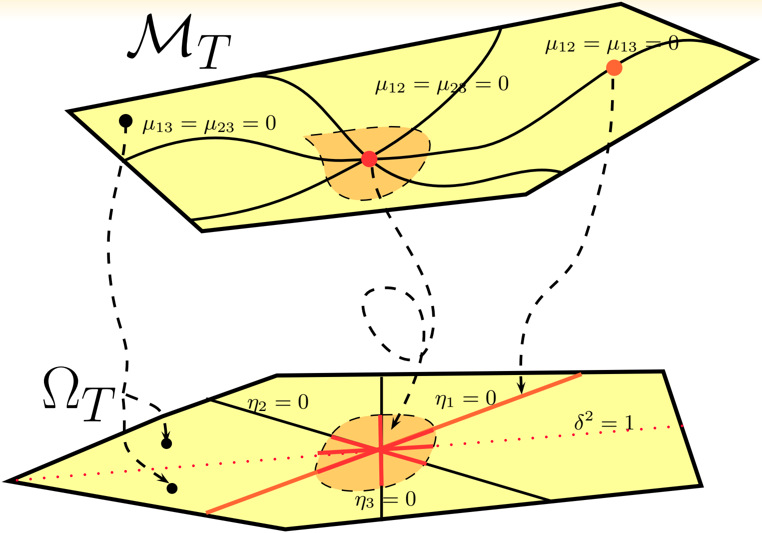
$$(t, a_1^{(0)}, a_1^{(1)}, b_1^{(0)}, b_1^{(1)}, c_1^{(0)}, c_1^{(1)}) \xleftrightarrow{1-1} (k_1, k_2, k_3, s, \bar{a}, \bar{b}, \bar{c})$$

The new parametrization

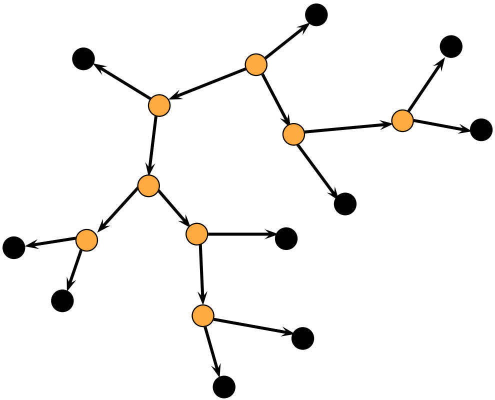
$$\mathcal{M}_T : \quad \begin{aligned} k_{12} &= \tfrac{1}{4}(1-s^2)\bar{a}\bar{b}, \\ k_{13} &= \tfrac{1}{4}(1-s^2)\bar{a}\bar{c}, \\ k_{23} &= \tfrac{1}{4}(1-s^2)\bar{b}\bar{c}, \\ k_{123} &= \tfrac{1}{4}(1-s^2)s\bar{a}\bar{b}\bar{c} \end{aligned}$$

◂ general formula

Application: Identifiability



Does it generalize?

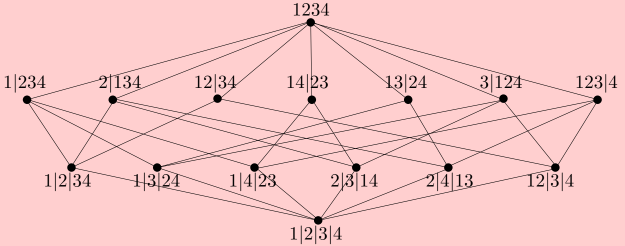


The general Markov model

- Let $T=(V,E)$ be a rooted tree with n leaves.
- for the root: $[\theta_0^{(r)}, \theta_1^{(r)}]$
- for each edge (u,v): $\theta_{i|j}^{(v)}$ for $i,j=0,1$ s.t. $\theta_{0|j}^{(v)}+\theta_{1|j}^{(v)}=1$
- $p_{\alpha_1\cdots\alpha_n}=\sum_{\beta\in\mathcal{H}(\alpha)}\theta_{\beta_r}^{(r)}\prod_{v\in V\setminus r}\theta_{\beta_v|\beta_{\mathrm{pa}(v)}}^{(v)}$ for all $\alpha\in\{0,1\}^n$, where $\mathcal{H}(\alpha)\subset\{0,1\}^V$ are sequences β such that $\beta_i=\alpha_i$ for all leaves i.
- e.g. for the tripod: $p_{ijk}=\theta_0^{(r)}\theta_{i|0}^{(1)}\theta_{j|0}^{(2)}\theta_{k|0}^{(3)}+\theta_1^{(r)}\theta_{i|1}^{(1)}\theta_{j|1}^{(2)}\theta_{k|1}^{(3)}$.

Poset of set partitions

- poset $\Pi(n)$ of all partitions of $[n] := \{1, \ldots, n\}$
  - $\pi = B_1|\cdots|B_r$ where $B_i \subseteq [n]$ s.t. $\bigcup_i B_i = [n]$ and $B_i \cap B_j = \emptyset$
- ordering
  - $\pi \leq \delta$ if every block of π is contained in one of the blocks of δ
  - e.g. $n = 4$ then $\hat{\mathbf{0}} = 1|2|3|4 < 1|4|23 < 14|23 < 1234 = \hat{\mathbf{1}}$



Moment-cumulant formula

Let $A \subseteq [n]$:

$$k_A = \sum_{\pi \in \Pi(A)} \mathfrak{m}(\pi, \hat{1}) \prod_{B \in \pi} \mu_B,$$

where $\mathfrak{m}(\pi, \hat{1}) = (-1)^{|\pi|-1}(|\pi| - 1)!$

If $n = 3$

- $k_{12} = \mu_{12} - \mu_1\mu_2$
- $k_{123} = \mu_{123} - \mu_1\mu_{23} - \mu_2\mu_{13} - \mu_3\mu_{12} + 2\mu_1\mu_2\mu_3$

L-cumulants

$$\ell_A = \sum_{\pi \in L(A)} \mathfrak{m}_L(\pi, \hat{1}) \prod_{B \in \pi} \mu_B,$$

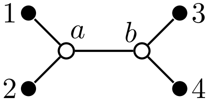
where $L(A)$ is any other poset of partitions with $\hat{0}$ and $\hat{1}$ as in $\Pi(|A|)$; and $\mathfrak{m}_L$ is the Möbius function on $L(A)$.

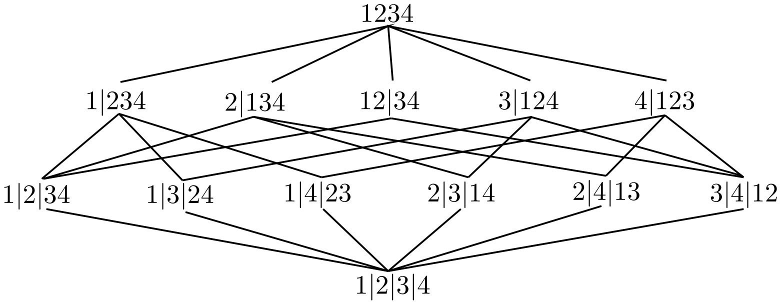
- Formally we need an inverse system $\mathbf{L} = (L(A))$ of partition lattices such that $L(A) \to L(B)$, for any two multisets $B \subseteq A$, is a restriction to B.

L-cumulants generalize...

- cumulants for all partitions of $[n]$
- free cumulants for non-crossing partitions
- Boolean cumulants for interval partitions
- central moments for one-cluster partitions

Tree partitions





Tree cumulants

- T tree with n leaves, $[n] = \{1, \ldots, n\}$ the set of leaves
- Π_T the poset of all the partitions of $[n]$ induced by removing inner nodes together with the Möbius function $\mathfrak{m}_T$

$$\ell^t_{1\cdots n} = \sum_{\pi \in \Pi_T([n])} \mathfrak{m}_T(\pi, \hat{1}) \prod_{B \in \pi} \mu_B$$

Reparameterization

- let $s_v = 1 - 2\mathbb{E}(Y_v)$ and $\eta_{uv} = \theta^{(v)}_{1|1} - \theta^{(v)}_{1|0}$ for all $(u,v) \in E$
- $\theta = (\theta^{(r)}_1, \theta^{(v)}_{1|0}, \theta^{(v)}_{1|1}) \overset{1-1}{\longleftrightarrow} \omega = (\delta_v, \eta_{uv})$

Theorem[Z.,Smith]

$$\ell^t_{1\cdots n} = \frac{1}{4}(1 - s_r^2) \prod_{v\in \mathrm{int}(V)} s_v^{\deg(v)-2} \prod_{e\in E} \eta_e$$

recall: tripod

Properties of L-cumulants (1)

Proposition [Z.] (semi-invariance)

Let $X' = X + a$. If $L(n)$ is such that $i|[n] \setminus i \in L(n)$ for every $i \in [n]$ then $\ell'_i = \ell_i + a_i$ and

$$\ell'_A = \ell_A \quad \text{for all } A \text{ such that } |A| \geq 2.$$

If $n = 3$ then there are four interval partitions: 123, $1|23$, $12|3$ and $1|2|3$ and hence $2|13 \notin L^{\text{int}}([3])$

$$\ell'_{123} = \ell_{123} + a_2 \ell_{13}.$$

Properties of L-cumulants (2)

Proposition [Z.] (multilinearity)

If $\mathbf{L}$ is such that $L(A) \simeq L(n)$ for every multiset A with n elements then $\mathrm{lcum}(X_1, \ldots, X_n)$ is multilinear.

Satisfied by non-crossing, interval and one-cluster partitions.

Properties of L-cumulants (3)

Proposition [Z.]

- For two random vectors we say they are L-independent if all mixed L-cumulants of X and Y vanish.
- If X and Y are L-independent then

$$\ell(X+Y)=\ell(X)+\ell(Y).$$

Conclusions and open questions

- L-cumulants generalize in an elegant way various probabilistic notions. However it is not entirely clear now to what extent this generalization is useful.
- L-cumulants can be used in various fields of mathematics. Is there a general theory of that?

Thank you!

The bibliography

- G. Pistone and H. Wynn, ***Cumulant varieties*, J. Symbolic Computation, 41(2), 2006.**
- T. P. Speed, ***Cumulants and partition lattices*, Austral. J. Statist, 25, 1983.**
- P. Zwiernik and J. Q. Smith, ***Tree-cumulants and the geometry of binary tree models*, to appear in *Bernoulli*.**
- P. Zwiernik, ***L-cumulants, L-cumulant embeddings and algebraic statistics*, 2010.**