# An algebraic characterisation of staged trees

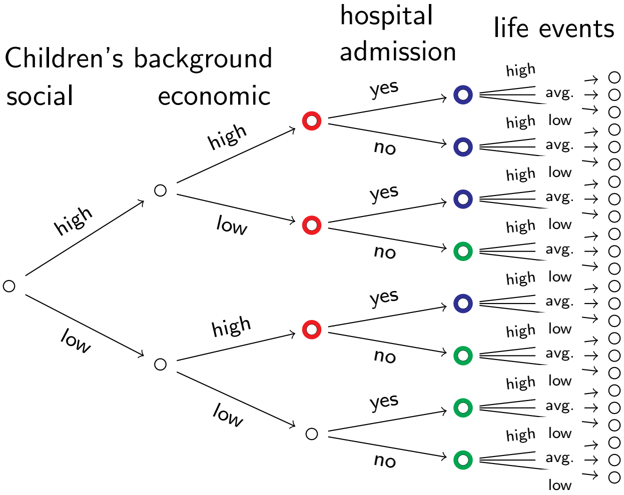
Christiane Görgen
and Jim Q. Smith

Genova – May 18, 2015

Warwick Statistics

This research is funded by EPSRC grant EP/L505110/1.

# The Christchurch Health & Development Study




Fergusson et al. (1986). Cowell and Smith (2015).

1. What do staged tree models look like?

2. How can we characterise a class of statistically equivalent staged trees?

3. Can we give a causal interpretation to the directionality of a tree graph?

# Ingredients

for a staged tree statistical model

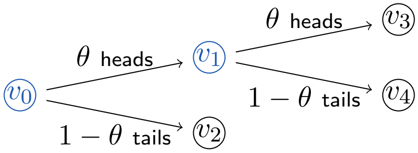
- a tree graph
- edge labels
- sum-to-$1$ conditions
- stage constraints

# Formalisation

A tree model (represented by a staged tree) is the class of all distributions which 'factorise according to the tree'.

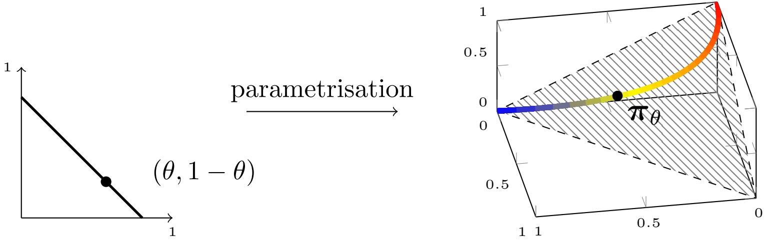
$$\mathbb{P} = \{\boldsymbol{\pi} \;\mid\; \boldsymbol{\pi}(\lambda) = \prod_{e \in E(\lambda)} \theta(e) \text{ for all atoms } \lambda\}$$

# Coin-toss example




$$\mathbb{P} = \left\{ \left(\theta^2, \theta(1-\theta), 1-\theta\right) \;\middle|\; \theta \in (0,1) \right\}$$

# Coin-toss model




$$\boldsymbol{\pi}_\theta = (\theta^2, \theta(1-\theta), 1-\theta)$$

# Usefulness of staged trees

Discrete graphical models with many advantages:

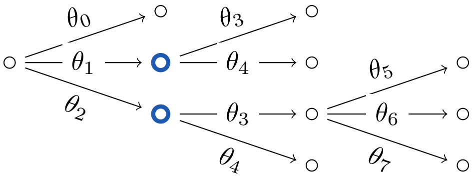
- highly expressive depiction of events
- easy to communicate
- can depict (context-specific) conditional independences
  $\rightsquigarrow$ more general than BNs
- useful for *asymmetric* modelling situations
- model selection and learning techniques $\rightsquigarrow$ chain event graphs

# When are two trees statistically equivalent?

represent the same model/ equivalent distributional assumptions
$\rightsquigarrow$ identified atoms have equal probabilities

- polynomial equivalence (same parametrisation)
- statistical equivalence

# Linking a tree graph and a polynomial

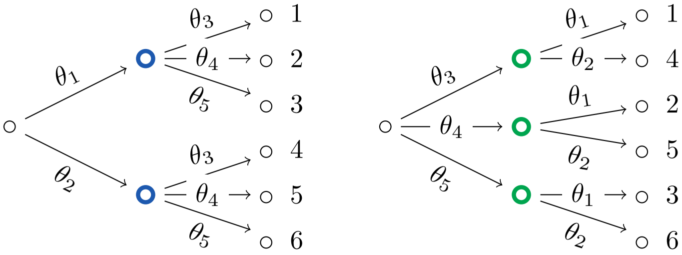



**Proposition** Every staged tree is in 1-to-1 correspondence with a nested polynomial

$$c_{\mathcal{T}}(\boldsymbol{\theta}) = \sum_{v_1 \in \mathrm{ch}(v_0)} \theta(v_0, v_1) \left( \sum_{v_2 \in \mathrm{ch}(v_1)} \theta(v_1, v_2) \left( \cdots \left( \sum_{v_k \in \mathrm{ch}(v_{k-1})} \theta(v_{k-1}, v_k) \right) \right) \right).$$

# Nesting a polynomial is not unique

Look for possible factorisations within the polynomial:




$$(\theta_1 + \theta_2)(\theta_3 + \theta_4 + \theta_5)$$

# Swap operator

distributive law on the polynomial $\longleftrightarrow$ local graph changes

changing the order of parameters changes the order of events and vice versa

**Lemma**
Polynomial equivalence $\Rightarrow$ statistical equivalence.

# Swaps and arc reversals

Consider

$$X_1 \longrightarrow X_2 \longrightarrow X_3$$

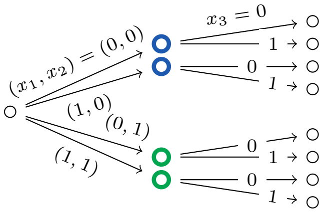
with probability mass function $p(x_1, x_2, x_3) = \theta(x_1, x_2)\theta(x_2, x_3)$ so

$$c(\boldsymbol{\theta}) = \sum_{x_1, x_2, x_3} \theta(x_1, x_2)\theta(x_2, x_3)$$

Nestings of $\sum_{x_1,x_2,x_3} \theta(x_1,x_2)\theta(x_2,x_3)$ include

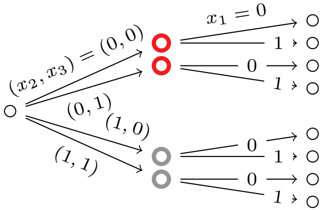
$$= \sum_{x_1,x_2} \theta(x_1,x_2) \left( \sum_{x_3} \theta(x_2,x_3) \right)$$

or $p_{12}(x_1,x_2)p_3(x_3|x_2)$




$$= \sum_{x_2,x_3} \theta(x_2,x_3) \left( \sum_{x_1} \theta(x_1,x_2) \right)$$

or $p_{23}(x_2,x_3)p_1(x_1|x_2)$



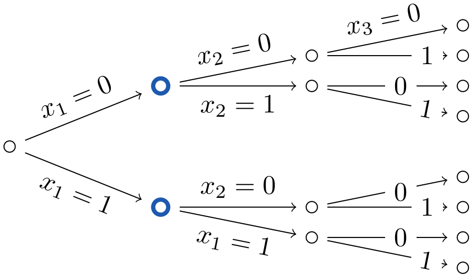

so $X_1 \longrightarrow X_2 \longrightarrow X_3$ becomes $X_1 \longleftarrow X_2 \longleftarrow X_3$.

# Independence and orders

Consider

$$X_1 \longrightarrow X_3 \longleftarrow X_2$$

with probability mass function
$p(x_1, x_2, x_3) = \theta(x_1)\theta(x_2)\theta(x_1, x_2, x_3)$ so

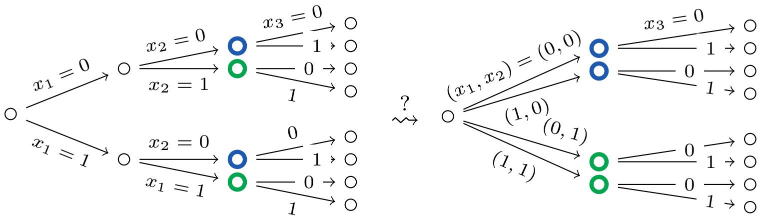


# Polynomial equivalence is not everything

Remember

$$X_1 \longrightarrow X_2 \longrightarrow X_3$$

so $p(x_1, x_2, x_3) = \theta(x_1)\theta(x_1, x_2)\theta(x_2, x_3)$ and



# Resize operator

substituting monomial terms $\longleftrightarrow$ shortening subtrees

algebraic operation: monomial or rational map on the polynomial

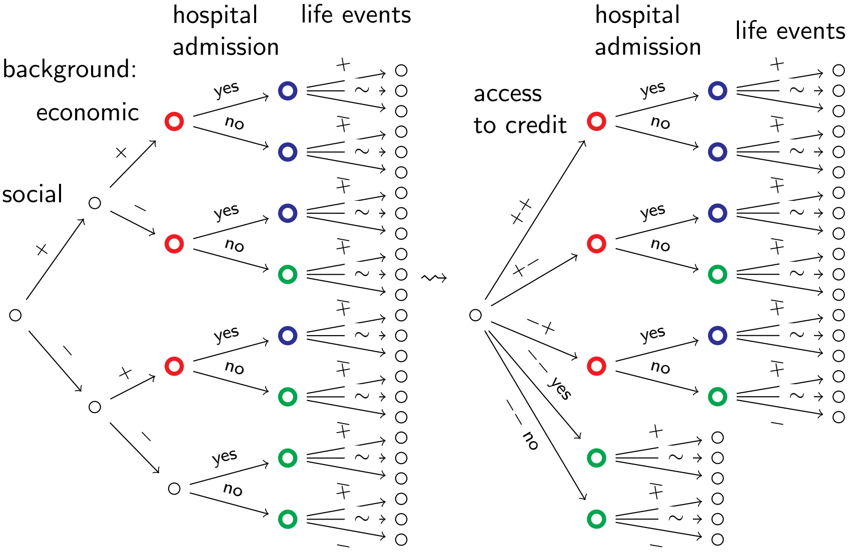
BN-analogue: clique-parametrisation in decomposable graphs

# Theorem

Between every two statistically equivalent staged trees there is a map which is a composition of swaps and resizes, and vice versa.

Görgen and Smith (2015). arXiv:1512.00209v2 [math.ST].

# CHDS again


hospital admission
life events
background:
economic
social
yes
no
access to credit
hospital admission
life events

# CHDS' statistical equivalence class

- create new variables: access to credit
- context-specific changes of the orders of events
- in all representations of the model: hospital admissions before life events $\rightsquigarrow$ putative causal interpretation

# Ongoing research

- CoCoA code which performs the swap/resize of a polynomial (with Eva Riccomagno and Anna Bigatti)
- causal discovery algorithm
- algebro-geometric characterisation of a staged tree model

Thanks for your attention!