The Geometry of Chain Event Graphs

Jim Smith and Christiane Görgen

University of Warwick

June 2015

The Plan of thisTalk

- An introduction to CEGs, staged trees and their relationship to BNs.
- How they can be used to describe a data set.
- What their polynomial structure looks like.
- Why the algebra gives extra insights about this model class.
- Equivalence classes and inferred causation.

I will suppress the mathematics here which will be given more formally in Christiane's poster.

Discrete Bayesian Networks for Multivariate Data

- BNs represent statistical relationships over product spaces elegantly, expressively & formally.
- Guide conjugate learning.& model selection.

However!

- BN specify dependences solely over a prespecified set of measurement variables.
- BN's not entirely natural when specifying relationships in terms of how processes might evolve.
- Sample space - often critical to estimation and selection issues - not depicted.
- Can only express certain types of probabilistic symmetry.

A BN (Barclay et al, 2012): Exploratory data analysis

$$\begin{tikzcd} & \text{Social Background} \arrow[dl] \arrow[d] \arrow[dr] & \\ \text{Economic Situation} \arrow[r] & \text{Family Life Events} \arrow[r] & \text{Hospital Admissions} \end{tikzcd}$$

- Study 1265 children over 5 years: HA 0 or at least 1, LE on 3 levels, Binary categories for ES & SB.
- Scored all 4 node BNs using standard Bayes Factor scoring rule.
- Best score amongst close competitors: where edges missing from ES$\rightarrow$LE, & one missing edge into HA. So given SB & LE, HA independent of ES.

.

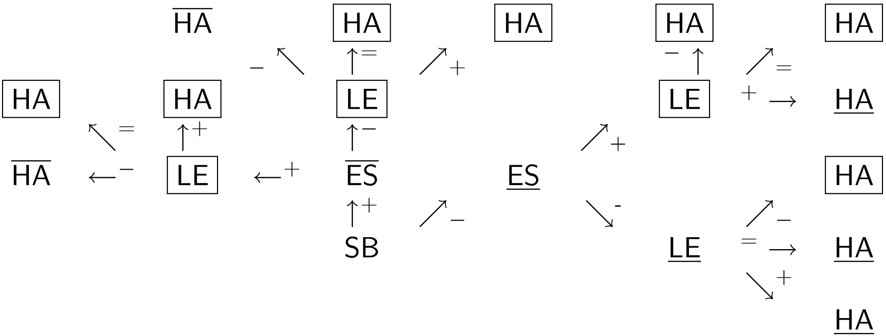
So why not use trees!

$$
\begin{array}{ccccccccccc}
 & & \text{HA} & & \text{HA} & & \text{HA} & & \text{HA} & & \text{HA} \\
 & & & - \nwarrow & \uparrow^{=} & \nearrow_{+} & & & - \uparrow & \nearrow_{=} & \\
\text{HA} & & \text{HA} & & \text{LE} & & & & \text{LE} & + \rightarrow & \text{HA} \\
 & \nwarrow^{=} & \uparrow^{+} & & \uparrow^{-} & & & \nearrow_{+} & & & \\
\text{HA} & \leftarrow^{-} & \text{LE} & \leftarrow^{+} & \text{ES} & & \text{ES} & & & & \text{HA} \\
 & & & & \uparrow^{+} & \nearrow_{-} & & \searrow^{\text{-}} & & \nearrow_{-} & \\
 & & & & \text{SB} & & & & \text{LE} & = \rightarrow & \text{HA} \\
 & & & & & & & & & \searrow^{+} & \\
 & & & & & & & & & & \text{HA}
\end{array}
$$

Can introduce conditional independence through equating edge probs associated with different nodes!!!!!

Example of staged tree (omitting leaves from HA)

.



- Colour partition $\left\{\text{SB}, \overline{\text{ES}}, \underline{\text{ES}}, \boxed{\text{LE}}, \underline{\text{LE}}, \overline{\text{HA}}, \boxed{\text{HA}}, \underline{\text{HA}}\right\}$: edge probs.

$$(\pi_{1s}, \pi_{2s}), (\pi_{1e}, \pi_{2e}), (\pi_{1\underline{e}}, \pi_{2\underline{e}}), (\pi_{1l}, \pi_{2l}, \pi_{3l}),$$
$$(\pi_{1\underline{l}}, \pi_{2\underline{l}}, \pi_{3\underline{l}}), (\pi_{1\overline{h}}, \pi_{2\overline{h}}), (\pi_{1h}, \pi_{2h}), (\pi_{1\underline{h}}, \pi_{2\underline{h}})$$

Example of staged tree (omitting leaves from HA)

.

- Colour partition, **stages** $\left\{\mathrm{SB}, \overline{\mathrm{ES}}, \underline{\mathrm{ES}}, \boxed{\mathrm{LE}}, \underline{\mathrm{LE}}, \overline{\mathrm{HA}}, \boxed{\mathrm{HA}}, \underline{\mathrm{HA}}\right\}$.
- **Positions** $\left\{\mathrm{SB}, \overline{\mathrm{ES}}, \underline{\mathrm{ES}}, \boxed{\mathrm{LE}_1}, \boxed{\mathrm{LE}_2}, \underline{\mathrm{LE}}, \overline{\mathrm{HA}}, \boxed{\mathrm{HA}}, \underline{\mathrm{HA}}\right\}$ - CEG nodes
- **Saturated model** with 24 atoms $=$ 23 dim. (atoms root -leaf paths).
- **CEG above** 18 edge probs (with 8 constraints) $=$ 10 dim.

$$\left\{\begin{array}{c} (\pi_{1s}, \pi_{2s}), (\pi_{1e}, \pi_{2e}), (\pi_{1\underline{e}}, \pi_{2\underline{e}}), (\pi_{1l}, \pi_{2l}, \pi_{3l}), \\ (\pi_{1\underline{l}}, \pi_{2\underline{l}}, \pi_{3\underline{l}}), (\pi_{1\overline{h}}, \pi_{2\overline{h}}), (\pi_{1h}, \pi_{2h}), (\pi_{1\underline{h}}, \pi_{2\underline{h}}) \end{array}\right\}$$

- **BN above** 32 edge probs (with 13 constraints) $=$ 19 dim.
- Smallest **independence model** $\amalg$SB,ES,LE,HA with 9 edge probs and 4 constraints $=$ 5 dim
- Staged tree MAP score was 80 **times better** than best BN.

Chain Event Graphs

- Simpler graph of staged tree showing sample space.
- Construction: Event tree $\longrightarrow$ Staged tree $\longrightarrow$ CEG
- Start with event tree & colour vertices - as illustrated above ($\longrightarrow$ staged tree).
- Identify positions which (with w_∞) form vertices of CEG.
- Construct CEG by inheriting edges from tree in obvious way $+$ attach all leaves to w_∞.

A top scoring CEG when HA the response.

$$
\begin{array}{ccccccc}
 & & & & \overline{\mathrm{HA}} & & \\
 & & & \text{-} \nearrow & & \searrow \searrow & \\
 & & \boxed{\mathrm{LE}} & \overset{=}{+} \rightrightarrows & \boxed{\mathrm{HA}} & \rightrightarrows & w_\infty \\
 & + \nearrow \; \nearrow_- & & - \nearrow \nearrow_= & \uparrow & & \upuparrows \\
\overline{\mathrm{ES}} & & \boxed{\mathrm{LE}} & + - & + & \rightarrow & \underline{\mathrm{HA}} \\
\uparrow^+ & & + \uparrow & & | & = \nearrow \nearrow^+ & \\
\mathrm{SB} & \_ \rightarrow & \underline{\mathrm{ES}} & \_ \rightarrow & \underline{\mathrm{LE}} & &
\end{array}
$$

- For SB$^+$,ES. has no impact on LE or HA .
- SB$^+$ & LE$^-$ lead to child most favorable HA.
- (SB$^+$ & LE$^{=,+}$) or (SB$^-$ & ES$^+$& LE$^{-,=}$) or (SB$^-$ & ES$^-$& LE$^+$) lead to moderate HA.
- (SB$^-$& ES$^-$& LE$^{=,+}$) or (SB$^-$& ES$^+$& LE$^+$) lead to worst HA.

Bayesian Inference on CEG's & Fast Learning

- Likelihood separates! so class of regular CEG's admits simple conjugate learning.
- Explicitly the likelihood under complete random sampling is given by

$$\begin{aligned} l(\boldsymbol{\pi}) &= \prod_{u \in U} l_u(\boldsymbol{\pi}_u) \\ l_u(\boldsymbol{\pi}_u) &= \prod_{i \in u} \pi_{i,u}^{x(i,u)} \end{aligned}$$

  where $x(i,u)$ # units entering stage u & proceeding along edge labelled (i,u), $\sum_i \pi_{u,i} = 1$
- Independent Dirichlet priors $D(\boldsymbol{\alpha}(u))$ on the vectors $\boldsymbol{\pi}_u$ leads to independent Dirichlet $D(\boldsymbol{\alpha}^*(u))$ posteriors where

$$\alpha^*(i,u) = \alpha(i,u) + x(i,u)$$

Score each CEG to find best explanation

- Score simple fn. of sampled data $\{x(i, u, \mathcal{C})\}$ counting units going from a stage then along edge in given CEG $\mathcal{C}$.
- Modular parameter priors over CEGs $\Rightarrow$ log marginal likelihood score *linear* in CEG stage scores. Select highest scoring $\mathcal{C}$
- For $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)$, let $s(\boldsymbol{\alpha}) = \log \Gamma(\sum_{i=1}^{k} \alpha_i)$ & $t(\boldsymbol{\alpha}) = \sum_{i=1}^{k} \log \Gamma(\alpha_i)$

$$\begin{aligned} \Psi(C) &= \log p(C) = \sum_{u \in C} \Psi_{u(c)} \\ \Psi_{u(c)} &= \sum s(\alpha(i,u)) - s(\alpha^*(i,u)) + t^*(\alpha(i,u)) - t(\alpha(i,u)) \end{aligned}$$

- e.g. MAP model selection/ NLP priors (Collazo & Smith, 2015) with D Prog (see Cowell & Smith,2014) or when nec. greedy search e.g. AHC $\rightarrow$ simple & fast over vast space of CEG's possible.
- Each CEG has an associated causal interpretation (see below).

Embellishing a CEG with probabilities

- Note that the positions in the same stage have the same associated edge probabilities.
- Probabilities of atoms calculated by multiplying up edge probabilities on each root to leaf path.

$$
\begin{array}{ccccccc}
 & & & & \overline{\mathrm{HA}} & & \\
 & & & \pi_{3l} \nearrow & & \pi_{1\overline{h}} \searrow \;\; \searrow \pi_{2\overline{h}} & \\
 & & \boxed{\mathrm{LE}} & \begin{array}{c}\pi_{2l} \\ \pi_{1l}\end{array} \rightrightarrows & \boxed{\mathrm{HA}} & \begin{array}{c}\pi_{1h} \\ \pi_{2h}\end{array} \rightrightarrows & w_{\infty} \\
 & \pi_{2e} \nearrow \;\; \nearrow \pi_{1e} & & \pi_{3l} \nearrow \;\; \nearrow \pi_{2l} & \pi_{3\underline{l}} \uparrow & & \pi_{1\underline{h}} \uparrow\uparrow \pi_{2\underline{h}} \\
\overline{\mathrm{ES}} & & \boxed{\mathrm{LE}} & \pi_{1l} - & + & \rightarrow & \underline{\mathrm{HA}} \\
\uparrow \pi_{2s} & & \pi_{2\underline{e}} \uparrow & & | & \pi_{2\underline{l}} \nearrow \;\; \nearrow \pi_{1\underline{l}} & \\
\mathrm{SB} & \pi_{1s} \rightarrow & \underline{\mathrm{ES}} & \pi_{1\underline{e}} \rightarrow & \underline{\mathrm{LE}} & &
\end{array}
$$

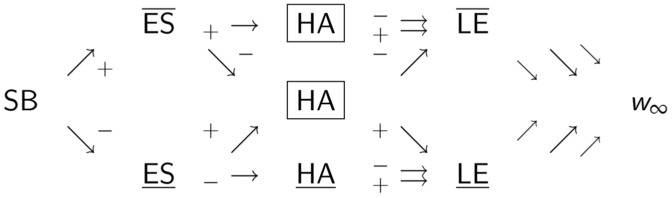
Atomic probs as monomials in primitive probs

$$
\begin{array}{ll}
p(\omega_1) = \pi_{2s}\pi_{2e}\pi_{3l}\pi_{2\overline{h}} & p(\omega_{13}) = \pi_{1s}\pi_{2\underline{e}}\pi_{3l}\pi_{2h} \\
p(\omega_2) = \pi_{2s}\pi_{2e}\pi_{3l}\pi_{1\overline{h}} & p(\omega_{14}) = \pi_{1s}\pi_{2\underline{e}}\pi_{3l}\pi_{1h} \\
p(\omega_3) = \pi_{2s}\pi_{2e}\pi_{2l}\pi_{2h} & p(\omega_{15}) = \pi_{1s}\pi_{2\underline{e}}\pi_{2l}\pi_{2h} \\
p(\omega_4) = \pi_{2s}\pi_{2e}\pi_{2l}\pi_{1h} & p(\omega_{16}) = \pi_{1s}\pi_{2\underline{e}}\pi_{2l}\pi_{1h} \\
p(\omega_5) = \pi_{2s}\pi_{2e}\pi_{1l}\pi_{2h} & p(\omega_{17}) = \pi_{1s}\pi_{2\underline{e}}\pi_{1l}\pi_{2\underline{h}} \\
p(\omega_6) = \pi_{2s}\pi_{2e}\pi_{1l}\pi_{1h} & p(\omega_{18}) = \pi_{1s}\pi_{2\underline{e}}\pi_{1l}\pi_{1\underline{h}} \\
p(\omega_7) = \pi_{2s}\pi_{1e}\pi_{3l}\pi_{2h} & p(\omega_{19}) = \pi_{1s}\pi_{1\underline{e}}\pi_{3\underline{l}}\pi_{2h} \\
p(\omega_8) = \pi_{2s}\pi_{1e}\pi_{3l}\pi_{1h} & p(\omega_{20}) = \pi_{1s}\pi_{1\underline{e}}\pi_{3\underline{l}}\pi_{1h} \\
p(\omega_9) = \pi_{2s}\pi_{1e}\pi_{2l}\pi_{2h} & p(\omega_{21}) = \pi_{1s}\pi_{1\underline{e}}\pi_{2\underline{l}}\pi_{2\underline{h}} \\
p(\omega_{10}) = \pi_{2s}\pi_{1e}\pi_{2l}\pi_{1h} & p(\omega_{22}) = \pi_{1s}\pi_{1\underline{e}}\pi_{2\underline{l}}\pi_{1\underline{h}} \\
p(\omega_{11}) = \pi_{2s}\pi_{1e}\pi_{1l}\pi_{2h} & p(\omega_{23}) = \pi_{1s}\pi_{1\underline{e}}\pi_{1\underline{l}}\pi_{2\underline{h}} \\
p(\omega_{12}) = \pi_{2s}\pi_{1e}\pi_{1l}\pi_{1h} & p(\omega_{24}) = \pi_{1s}\pi_{1\underline{e}}\pi_{1\underline{l}}\pi_{1\underline{h}}
\end{array}
$$

- Because based on BN monomials are all of same degree (a property not required for CEGs). But with less symmetry in indeterminates!.

Example CHIDS a different CEG

A best model identified through Dynamic Programming allowing changed response variable.



- This model sees *life events as a result of poor child health*.
- Increased incidents of hospital admissions relates only to poverty (2 categories).
- High life events unaffected by Hospital Admissions except that when exactly one of SB or ES is low then poor child health can shift into lower life event category.

New atomic probabilities

Now have stages $\{SB, \overline{ES}, \underline{ES}, \boxed{HA}, \underline{HA}, \underline{LE}, \overline{LE}\}$ with 16 parameters and 7 constraints = 9 dim space

$$
\begin{array}{ll}
p(\omega_1) = \pi_{2s}\pi_{2e}\pi_{2h}\pi_{3\bar{l}} & p(\omega_{13}) = \pi_{1s}\pi_{2\underline{e}}\pi_{2h}\pi_{3l} \\
p(\omega_2) = \pi_{2s}\pi_{2e}\pi_{1h}\pi_{3\bar{l}} & p(\omega_{14}) = \pi_{1s}\pi_{2\underline{e}}\pi_{1h}\pi_{3l} \\
p(\omega_3) = \pi_{2s}\pi_{2e}\pi_{2h}\pi_{2l} & p(\omega_{15}) = \pi_{1s}\pi_{2\underline{e}}\pi_{2h}\pi_{2l} \\
p(\omega_4) = \pi_{2s}\pi_{2e}\pi_{1h}\pi_{2l} & p(\omega_{16}) = \pi_{1s}\pi_{2\underline{e}}\pi_{1h}\pi_{2l} \\
p(\omega_5) = \pi_{2s}\pi_{2e}\pi_{2h}\pi_{1l} & p(\omega_{17}) = \pi_{1s}\pi_{2\underline{e}}\pi_{2\underline{h}}\pi_{1l} \\
p(\omega_6) = \pi_{2s}\pi_{2e}\pi_{1h}\pi_{1l} & p(\omega_{18}) = \pi_{1s}\pi_{2\underline{e}}\pi_{1\underline{h}}\pi_{1l} \\
p(\omega_7) = \pi_{2s}\pi_{1e}\pi_{2h}\pi_{3l} & p(\omega_{19}) = \pi_{1s}\pi_{1\underline{e}}\pi_{2\underline{h}}\pi_{3\underline{l}} \\
p(\omega_8) = \pi_{2s}\pi_{1e}\pi_{1h}\pi_{3l} & p(\omega_{20}) = \pi_{1s}\pi_{1\underline{e}}\pi_{1\underline{h}}\pi_{3\underline{l}} \\
p(\omega_9) = \pi_{2s}\pi_{1e}\pi_{2h}\pi_{2l} & p(\omega_{21}) = \pi_{1s}\pi_{1\underline{e}}\pi_{2\underline{h}}\pi_{2\underline{l}} \\
p(\omega_{10}) = \pi_{2s}\pi_{1e}\pi_{1h}\pi_{2l} & p(\omega_{22}) = \pi_{1s}\pi_{1\underline{e}}\pi_{1\underline{h}}\pi_{2\underline{l}} \\
p(\omega_{11}) = \pi_{2s}\pi_{1e}\pi_{2h}\pi_{1l} & p(\omega_{23}) = \pi_{1s}\pi_{1\underline{e}}\pi_{2\underline{h}}\pi_{1\underline{l}} \\
p(\omega_{12}) = \pi_{2s}\pi_{1e}\pi_{1h}\pi_{1l} & p(\omega_{24}) = \pi_{1s}\pi_{1\underline{e}}\pi_{1\underline{h}}\pi_{1\underline{l}}
\end{array}
$$

Interpretation & equivalent models Görgen & Smith(2015)

- Likelihoods of 2 *statistically equivalent (se)* CEGs will always be the same: regardless of data.
- To interpret results of search need to determine *what topological features are shared* across equivalence class & which differ.
- In above example best CEG has HA causing LE: but is this true for all se CEGs - or is there an equivalent model which appear to suggest LE causes HA? If so then clearly cannot convincingly propose HA causes LE!!!
- All good scoring methods will score these models the same. But often not able to search whole of space so not score all equivalence class.
- Two discrete BNs are se iff they the same essential *graph* (or pattern).
- However need algebraic characterization (not graphical) for CEGs!!

Determining equivalent statistical models Görgen and Smith(2015)

Definition

The *interpolating polynomial* $C(\boldsymbol{\pi})$ of a CEG G whose root to sink paths/atoms $\omega \in \Omega$ have associated probabilities monomials $\lambda_{\omega}^{G}(\boldsymbol{\pi})$ in $\boldsymbol{\pi}(G)$ the vector of all edge probabilities in G is given by

$$C^{G}(\boldsymbol{\pi}) \triangleq \sum_{\omega \in \Omega} c_{\omega} \lambda_{\omega}^{G}(\boldsymbol{\pi})$$

where $\{c_{\omega} : \omega \in \Omega\}$ are indicators on the atoms, not depending on G.

Theorem

If $C^{G_1}(\boldsymbol{\pi}) = C^{G_2}(\boldsymbol{\pi})$ then the CEGs G_1, G_2 are statistically equivalent .

Can ignore sum to one conditions on $\boldsymbol{\pi}(G)$. Statistical equivalence corresponds to existence of maps between interpolating polynomials: characterising $\sim$ for many classes of CEG - see Görgen & Smith (2015).

Orbiting an equivalence class with swaps and contractions

Definition

Say G_1 & G_2 are *polynomially equivalent* iff $C^{G_1}(\boldsymbol{\pi}) = C^{G_2}(\boldsymbol{\pi})$.

- By last theorem $C(\boldsymbol{\pi})$ becomes *label* for a particular probability model associated with many topologically different CEGs just as topology of a BN embeds many equivalent factorizations under different partial orders.

Theorem

Two CEGs G_1, G_2 are polynomially equivalent iff G_2 can be obtained from G_1 through sequence of swap operations.

- Formal definition of swap in Christaine's poster & Görgen & Smith (2015).

Example of Swap

$$\begin{array}{ccccc} & \curvearrowright & & & \\ b & \rightrightarrows & a & \rightrightarrows & e \\ \downarrow & \searrow & & & \\ d & & c & & \end{array} \quad \text{to} \quad \begin{array}{ccccc} & & & \curvearrowright & \\ a & \rightrightarrows & b & \rightrightarrows & e \\ \downarrow & \searrow & & & \\ d & & c & & \end{array}$$

- "Arc reversals" allow us to transverse set of *all* equivalent BNs.
- Swaps do the same for polynomial equivalent models! But now a collection of matrix operations.

Two statistically equivalent CEGs need not be polynomially equivalent.

$$\begin{tikzcd} a \arrow[r, "\pi_2"] \arrow[d, "\pi_1"] & b \arrow[r, "\phi_2"] \arrow[dr, "\phi_2"] & \omega_3 \\ \omega_1 & \mathbf{G}_1 & \omega_2 \end{tikzcd} \quad \text{to} \quad \begin{tikzcd} a \arrow[rr, "\pi_3'"] \arrow[d, "\pi_1"] \arrow[drr, "\pi_2'"] & & \omega_3 \\ \omega_1 & \mathbf{G}_2 & \omega_2 \end{tikzcd}$$

- Here G_1 statistically equivalent to G_2 - both saturated model on $\{\omega_1, \omega_2, \omega_3\}$. But

$$\begin{aligned} C^{G_1}(\boldsymbol{\pi}) &= c_1\pi_1 + c_2\pi_2\phi_2 + c_3\pi_2\phi_3 \\ C^{G_2}(\boldsymbol{\pi}') &= c_1\pi_1 + c_2\pi_2' + c_3\pi_3' \end{aligned}$$

so not polynomially equivalent!!!!

- Need additional local operation called *resize* to traverse whole space for general CEGs.

Return to the CHIDs example

Question In our best scoring model is there a statistically equivalent model that has a CEG representation with LR before HA?
If so then there is no reason to conjecture that Hospital Admissions cause Life Events and not vice versa.

- Exhaustive search demonstrates that - at least over those models that retain SB, ES, HA, LE strata all se models have $HA \prec LE$.
- More elegantly the same result can be shown by demonstrating that no sequence of contraction/expansion or swaps allows us to have $HA \prec LE$ within this class.

Conclusions

- Usefulness of CEGs in **biology, social processes, health & forensic science** now established.
- Like a BN, a CEG embeds certain causal conjectures that can be tested.
- Like a BN, a CEG has associated vector of polynomials $\Rightarrow$ properties of a CEG usefully formalised & examined using techniques of algebraic geometry - see Christiane's poster.
- In particular **computer algebra** can be used to determine when two CEGs are **statistically indistinguishable**, explore the **sensitivity** of a given model & **proximity** of models within the class & examine **identifiability** of class & **properties of estimators**.
- Discovering **causal** explanations behind a CEG, consistent across the discovered equivalence classes are especially useful in applications.

THANK YOU FOR YOUR ATTENTION!!

Selected References of the authors

Görgen, C. & Smith, J.Q. (2015) "Equivalence Classes of Chain Event Graphs" (in prep.)
Collazo, R.A. & Smith, J.Q.(2015) "A new family of Non-local Priors for Chain Event Graph model selection" CRiSM Res.Rep. 15 -02 (submitted)
Görgen, C. Leonelli, M. & Smith, J.Q. (2015) "A Differential Approach for Staged Trees" Proceeding of ESQAR conference July '15
Thwaites P.A.& Smith J.Q. (2015) "A Separation Theorem for Chain Event Graphs (submitted)
Cowell, R.G.& Smith, J.Q. (2014) "Causal discovery through MAP selection of stratified chain event graphs" Electronic J of Statistics vol.8, 965 - 997
Barclay, L.M., Hutton, J.L.& Smith, J.Q. (2014) "Chain Event Graphs for Informed Missingness" Bayesian Analysis, 9,1, 53-76
Barclay, L.M. , Hutton, J.L. & Smith, J.Q.(2013) "Refining a Bayesian Network using a Chain Event Graph" International J. of Approximate Reasoning 54, 1300-1309.

Selected References of the authors

Freeman, G. & Smith, J.Q. (2011) "Dynamic Staged Trees for Discrete Multivariate Time Series: Forecasting, Model Selection & Causal Analysis", Bayesian Analysis,6,2, 279 - 306

Freeman, G. & Smith, J.Q. (2011a) " Bayesian MAP Selection of Chain Event graphs" J. Multivariate Analysis, 102, 1152 -1165

Thwaites, P. Smith, J.Q. and Riccomagno, E. (2010) "Causal Analysis with Chain Event Graphs" Artificial Intelligence, 174, 889–909

Riccomagno, E.& Smith, J.Q. (2009) "The Geometry of Causal Probability Trees that are Algebraically Constrained" in "Optimal Design & Related Areas in Optimization and Statistics" Eds L. Pronzato & A.Zhigljavsky, Springer 131-152

Smith, J.Q. & Anderson P.E. (2008) "Conditional independence & Chain Event Graphs" Artificial Intelligence, 172, 1, 42 - 68

Riccomagno, E.M. & Smith, J.Q. (2004) "Identifying a cause in models which are not simple Bayesian networks" Proceedings of IMPU, Perugia July 04, 1315-22