Conditional independence and Markov properties for undirected graphs

Aarhus University, Fall 2003, Lecture 1

Steffen L. Lauritzen, Aalborg University

Conditional independence

Random variables X and Y are *conditionally independent* given the random variable Z if

$$\mathcal{L}(X \mid Y, Z) = \mathcal{L}(X \mid Z).$$

We then write $X \perp\!\!\!\perp Y \mid Z$ (or $X \perp\!\!\!\perp_P Y \mid Z$)

Intuitively:

Knowing Z renders Y *irrelevant* for predicting X.

Factorisation of densities:

$$\begin{aligned} X \perp\!\!\!\perp Y \mid Z \quad &\iff \quad f(x,y,z)f(z) = f(x,z)f(y,z) \\ &\iff \quad \exists a, b : f(x,y,z) = a(x,z)b(y,z). \end{aligned}$$

Fundamental properties

For random variables X, Y, Z, and W it holds

(C1) if $X \perp\!\!\!\perp Y \,|\, Z$ then $Y \perp\!\!\!\perp X \,|\, Z$;

(C2) if $X \perp\!\!\!\perp Y \,|\, Z$ and $U = g(Y)$, then $X \perp\!\!\!\perp U \,|\, Z$;

(C3) if $X \perp\!\!\!\perp Y \,|\, Z$ and $U = g(Y)$, then $X \perp\!\!\!\perp Y \,|\, (Z, U)$;

(C4) if $X \perp\!\!\!\perp Y \,|\, Z$ and $X \perp\!\!\!\perp W \,|\, (Y, Z)$, then $X \perp\!\!\!\perp (Y, W) \,|\, Z$;

If density w.r.t. product measure $f(x, y, z, w) > 0$ also

(C5) if $X \perp\!\!\!\perp Y \,|\, (Z, W)$ and $X \perp\!\!\!\perp Z \,|\, (Y, W)$ then $X \perp\!\!\!\perp (Y, Z) \,|\, W$.

Additional note on (C5)

$f(x, y, z, w) > 0$ is *not necessary* for (C5). Enough e.g. that $f(y, z, w) > 0$ or $f(x, z, w) > 0$; see proof in Lauritzen (1996).

In discrete and finite case it is even enough that for all w with $f(w) > 0$ the bipartite graphs $\mathcal{G}_w = (\mathcal{Y} \cup \mathcal{Z}, E_w)$ defined by

$$y \sim_w z \iff f(y, z, w) > 0,$$

are all connected.

Alternatively with X replacing Y.

Is there a simple necessary and sufficient condition?

Graphoid axioms

Ternary relation $\perp_\sigma$ is *graphoid* if for all disjoint subsets A, B, C, and D of V:

(S1) if $A \perp_\sigma B \,|\, C$ then $B \perp_\sigma A \,|\, C$;

(S2) if $A \perp_\sigma B \,|\, C$ and $D \subseteq B$, then $A \perp_\sigma D \,|\, C$;

(S3) if $A \perp_\sigma B \,|\, C$ and $D \subseteq B$, then $A \perp_\sigma B \,|\, (C \cup D)$;

(S4) if $A \perp_\sigma B \,|\, C$ and $A \perp_\sigma D \,|\, (B \cup C)$, then $A \perp_\sigma (B \cup D) \,|\, C$;

(S5) if $A \perp_\sigma B \,|\, (C \cup D)$ and $A \perp_\sigma C \,|\, (B \cup D)$ then $A \perp_\sigma (B \cup C) \,|\, D$.

Semigraphoid if only (S1)–(S4) holds.

Separation in undirected graphs

Let $\mathcal{G} = (V, E)$ be finite and simple undirected graph (no self-loops, no multiple edges).

For subsets A, B, S of V, let $A \perp_{\mathcal{G}} B \mid S$ denote that S *separates* A *from* B *in* $\mathcal{G}$, i.e. that all paths from A to B intersect S.

Fact: *The relation* $\perp_{\mathcal{G}}$ *on subsets of* V *is a graphoid.*

This fact is the reason for choosing the name 'graphoid' for such separation relations.

Probabilistic semigraphoids

V finite set, $X = (X_v, v \in V)$ random variables.

For $A \subseteq V$, let $X_A = (X_v, v \in A)$.

Let $\mathcal{X}_v$ denote state space of X_v.

Similarly $x_A = (x_v, v \in A) \in \mathcal{X}_A = \times_{v \in A} \mathcal{X}_v$.

Abbreviate: $A \perp\!\!\!\perp B \,|\, S \iff X_A \perp\!\!\!\perp X_B \,|\, X_S$.

Then basic properties of conditional independence imply:

The relation $\perp\!\!\!\perp$ *on subsets of* V *is a semigraphoid.*

If $f(x) > 0$ *for all* x*,* $\perp\!\!\!\perp$ *is also a graphoid.*

Not all semigraphoids are probabilistically representable.

Markov properties for semigraphoids

$\mathcal{G} = (V, E)$ simple undirected graph; $\perp_\sigma$ (semi)graphoid relation. Say $\perp_\sigma$ satisfies

(P) *the pairwise Markov property* if

$$\alpha \not\sim \beta \implies \alpha \perp_\sigma \beta \,|\, V \setminus \{\alpha, \beta\};$$

(L) *the local Markov property* if

$$\forall \alpha \in V : \alpha \perp_\sigma V \setminus \mathrm{cl}(\alpha) \,|\, \mathrm{bd}(\alpha);$$

(G) *the global Markov property* if

$$A \perp_\mathcal{G} B \,|\, S \implies A \perp_\sigma B \,|\, S.$$

Structural relations among Markov properties

For any semigraphoid it holds that

$$\text{(G)} \implies \text{(L)} \implies \text{(P)}$$

If $\perp_\sigma$ satisfies graphoid axioms it further holds that

$$\text{(P)} \implies \text{(G)}$$

so that *in the graphoid case*

$$\text{(G)} \iff \text{(L)} \iff \text{(P)}.$$

The latter holds in particular for $\perp\!\!\!\perp$, when $f(x) > 0$.

Factorisation and Markov properties

Assume density f w.r.t. product measure on $\mathcal{X}$.

For $a \subseteq V$, $\psi_a(x)$ depends on x_a only. Distribution of X *factorizes w.r.t.* $\mathcal{G}$ or satisfies (F) if

$$f(x) = \prod_{a \in \mathcal{A}} \psi_a(x)$$

where $\mathcal{A}$ are *complete* subsets of $\mathcal{G}$. *It then holds that*

$$\text{(F)} \implies \text{(G)}$$

and further: *If $f(x) > 0$ for all x*, (P) $\implies$ (F), so then

$$\text{(F)} \iff \text{(G)} \iff \text{(L)} \iff \text{(P)}.$$