Bayesian Networks 4
I-Equivalence, Distributions to Graphs

1

Soundness and Completeness

For d-separation, we would like:

1) Soundness: i.e. D-separation in Bayesian Network $\mathcal{G}$ guarantees conditional independence in distribution *P*
2) Completeness: i.e. D-separation in Bayesian Network $\mathcal{G}$ detects all possible independences in distribution *P*

Let's find out if both are true.

2

Soundness and Completeness

For d-separation, we would like:

1) Soundness: i.e. D-separation in Bayesian Network $\mathcal{G}$ guarantees conditional independence in distribution *P*

True: see proof in Section 4.5.1.1 ✓

3

Soundness and Completeness

For d-separation, we would like:

1) Soundness: i.e. D-separation in Bayesian Network $\mathcal{G}$ guarantees conditional independence in distribution *P* ✓

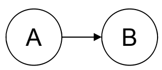
2) Completeness: i.e. D-separation in Bayesian Network $\mathcal{G}$ detects all possible independences in distribution *P**

**Not completely true*

4

D-separation

- Some independencies cannot be read off from the graph structure
- There may be additional (conditional) independencies in the graph that are not detected by d-separation
- (Example below): *A* is really independent of *B* after inspecting the Conditional Probability Tables



A	B	P(B|A)
false	false	0.4
false	true	0.6
true	false	0.4
true	true	0.6

5

Soundness and Completeness

For d-separation, we would like:

1) Soundness: i.e. D-separation in Bayesian Network $\mathcal{G}$ guarantees conditional independence in distribution *P*

2) Completeness: i.e. D-separation in Bayesian Network $\mathcal{G}$ detects all possible independences in distribution *P**

**True for the most part. Cases that violate these are rare and slight perturbations to the CPDs will eliminate these cases*

6

I-Equivalence

Very different BN structures can actually encode the same set of conditional independence assertions eg. the three structures below encode $(X \perp Y \mid \mathbf{Z})$:

$X \rightarrow Z \rightarrow Y$

$X \leftarrow Z \rightarrow Y$

$X \leftarrow Z \leftarrow Y$

Two graph structures $\mathcal{K}_1$ and $\mathcal{K}_2$ over $\mathcal{X}$ are I-equivalent if $\mathcal{I}(\mathcal{K}_1) = \mathcal{I}(\mathcal{K}_2)$.

7

I-Equivalence

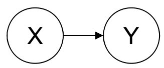
I-equivalence of two graphs implies:

- Any distribution *P* that can be factorized over one of these graphs can be factorized over other
- *P* can be associated with either graph

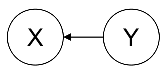
8

I-Equivalence

- Suppose we know that *X* and *Y* are correlated in the distrubtion *P(X,Y)*
- We don’t know if the correct structure is:



Or

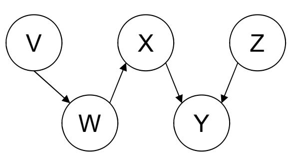


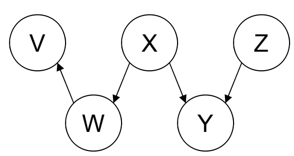
This has big implications for inferring causality! We’ll cover this later in the course if we have time

9

I-Equivalence

The skeleton of a Bayesian network graph $\mathcal{G}$ over $\mathcal{X}$ is an undirected graph over $\mathcal{X}$ that contains an edge $\{X,Y\}$ for every edge (X,Y) in $\mathcal{G}$





These two BNs have the same skeleton

10

I-Equivalence

- If two networks have a common skeleton, then the set of trails between two variables is the same in both networks
- But...having the same trails is not enough for I-equivalence eg.

$X \rightarrow Z \rightarrow Y$

$X \leftarrow Z \leftarrow Y$

$X \leftarrow Z \rightarrow Y$

Not I-equivalent to: $X \rightarrow Z \leftarrow Y$

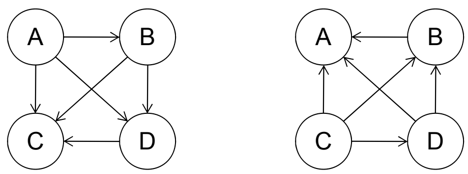
I-Equivalence

Theorem 3.7: Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two graphs over $\mathcal{X}$. If $\mathcal{G}_1$ and $\mathcal{G}_2$ have the same skeleton and the same set of v-structures then they are I-equivalent.

I-Equivalence

But there are graphs that are I-equivalent but do not have the same set of v-structures

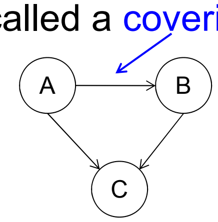
- eg. two complete (fully-connected) graphs have the same skeleton but not the same v-structures.



Can we provide a stronger condition that corresponds to I-Equivalence?

I-Equivalence

A v-structure $X \rightarrow Z \leftarrow Y$ is an immorality if there is no direct edge between X and Y. If there is such an edge, it is called a covering edge for the v-structure.



Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two graphs over $\mathcal{X}$. Then $\mathcal{G}_1$ and $\mathcal{G}_2$ have the same skeleton and the same set of immoralities if and only if they are I-equivalent

Distributions to Graphs

15

Distributions to Graphs

Given a distribution P, to what extent can we construct a graph $\mathcal{G}$ whose independencies reflect those of P?

16

Distributions to graphs

A graph $\mathcal{K}$ is a perfect map (P-map) for a set of independencies $\mathcal{I}$ if we have that $\mathcal{I}(\mathcal{K}) = \mathcal{I}$. We say that $\mathcal{K}$ is a perfect map for P if $\mathcal{I}(\mathcal{K}) = \mathcal{I}(P)$.

17

Distributions to Graphs

Does every distribution have a perfect map?
No…2 common counterexamples

1. Regularity in the parameterization of the distribution (eg. XOR relationships) that cannot be captured in the graph structure
2. Independence assumptions imposed by the structure of BNs is not appropriate

18

Distributions to Graphs

Counterexample of type 1:

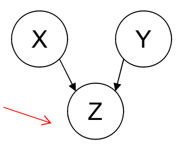
$$P(x,y,z)=\begin{cases}1/12 & x\oplus y\oplus z = \text{false}\\ 1/6 & x\oplus y\oplus z = \text{true}\end{cases}$$

Exclusive (XOR)

x	y	z	x⊕y⊕z	P(x,y,z)
0	0	0	false	1/12
0	0	1	true	1/6
0	1	0	true	1/6
0	1	1	false	1/12
1	0	0	true	1/6
1	0	1	false	1/12
1	1	0	false	1/12
1	1	1	true	1/6

$(X \perp Y) \in I(P)$ but
$(X \perp Z \mid Y) \notin I(P)$ and
$(Y \perp Z \mid X) \notin I(P)$

One possible minimal I-map:



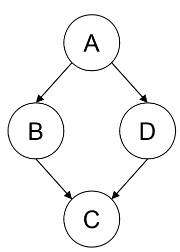
XOR CPD in Z

But it is not a perfect map since $(X \perp Z) \in I(P)$ which cannot be "read off" the graph by d-separation

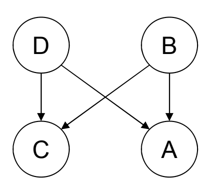
19

Distributions to Graphs

Counterexample of type 2:
Suppose we have $(A \perp C \mid \{B,D\}) \in I(P)$ and
$(B \perp D \mid \{A,C\}) \in I(P)$
Can we draw a P-map with just these independencies? No



This say $(B \perp D \mid A) \in I(P)$



This say $(B \perp D) \in I(P)$

Can't express these independencies with a BN. You need an undirected graphical model

20