3.5 Interventions

We can modify an SEM by setting e.g. $Z_k = a$. This is called a <u>perfect intervention</u>.
The new SEM determines a new distribution for Z.
Expectations or probabilities under the new distribution are written by adding "$do(Z_k = a)$"
e.g. $\mathbb{E}(Z_j \mid do(Z_k = a))$. Not necessarily true that $\mathbb{E}(Z_j \mid do(Z_k = a)) = \mathbb{E}(Z_j \mid Z_k = a)$.

<u>Example cont</u> After intervention $do(Z_2 = 1)$ (everyone is forced to go to the catch up lectures), we have a new SEM:

$$Z_3 = \varepsilon_3 \sim \mathrm{Bern}\left(\tfrac{1}{4}\right) \qquad Z_2 = 1$$

$$Z_1 = \mathbb{1}_{\{\varepsilon_1(1+Z_3) > \frac{1}{2}\}} \qquad \varepsilon_1 \sim U[0,1]$$

Thus $\mathbb{P}(Z_1 = 1 \mid do(Z_2 = 1)) = \frac{3}{4} \times \frac{1}{4} + \frac{1}{2} \times \frac{3}{4} = \frac{9}{16}$

$$\mathbb{P}(Z_1 = 1 \mid Z_2 = 1) = \sum_{j \in \{0,1\}} \mathbb{P}(Z_1 = 1 \mid Z_2 = 1, Z_3 = j)\, \mathbb{P}(Z_3 = j \mid Z_2 = 1)$$

$$= \sum_{j \in \{0,1\}} \mathbb{P}(Z_1 = 1 \mid Z_2 = 1, Z_3 = j) \frac{\mathbb{P}(Z_2 = 1 \mid Z_3 = j)\, \mathbb{P}(Z_3 = j)}{\mathbb{P}(Z_2 = 1)}$$

$$= \frac{16}{9}\left(\frac{3}{4} \times \frac{3}{4} \times \frac{1}{4} + \frac{1}{2} \times \frac{1}{2} \times \frac{3}{4}\right) = \frac{7}{12} \neq \frac{9}{16}$$

3.6 Markov properties on DAGs

Let P be the distribution of Z and suppose it has density f.

<u>Def^n</u> Given a DAG $\mathcal{G}$, say P satisfies the

(i) the <u>Markov factorisation property</u> w.r.t. $\mathcal{G}$ if

$$f(z_1, \dots, z_p) = \prod_{k=1}^{p} f(z_k \mid z_{pa(k)})$$

(ii) the <u>global Markov property</u> w.r.t. $\mathcal{G}$ if $\forall$ disjoint $A, B, S \subseteq \{1, \dots, p\}$ $(A, B \neq \emptyset)$

$$A, B \text{ d-separated by } S \Rightarrow Z_A \perp\!\!\!\perp Z_B \mid Z_S \quad (\text{if } S = \emptyset \text{ then } Z_A \perp\!\!\!\perp Z_B)$$

<u>Theorem 32</u> If P has a density f w.r.t. a product measure then (i) & (ii) above are equivalent.

We will often use "Markov" to mean "global Markov".

Prop 33 Let P be the distribution given by an SEM with DAG $\mathcal{G}$. Then P satisfies the Markov factorisation property w.r.t. $\mathcal{G}$.

Proof: Let π be a topological ordering of $\mathcal{G}$ and write $\tau = \pi^{-1}$. We can write

$$f(z_1, \dots, z_p) = f(z_{\tau(1)})\, f(z_{\tau(2)} \mid z_{\tau(1)})\, f(z_{\tau(3)} \mid z_{\tau(2)}, z_{\tau(1)}) \cdots f(z_{\tau(p)} \mid z_{\tau(p-1)}, \dots, z_{\tau(1)})$$

Now $pa(\tau(k)) \subseteq \{\tau(1), \dots, \tau(k-1)\}$. Since $Z_{\tau(k)}$ is a function of $Z_{pa(\tau(k))}$ and independent $\varepsilon_{\tau(k)}$, we know that $f(z_{\tau(k)} \mid z_{\tau(k-1)}, \dots, z_{\tau(1)}) = f(z_{\tau(k)} \mid z_{pa(\tau(k))})$.
Now substitute into eqn above. □

3.7 Causal structure learning

Given P, how can we find the DAG of the SEM that generated it?

3.7.1 Three obstacles

Causal minimality

If P is generated by an SEM with DAG $\mathcal{G}$, then P is Markov w.r.t $\mathcal{G}$.
Conversely, if P is Markov w.r.t. DAG $\mathcal{G}$, then is an SEM with DAG $\mathcal{G}$ that could have generated P.
Indeed take

$$Z_j = h_j(Z_{pa(j)}, \varepsilon_j) \quad \text{where } h_j(\tilde{z}_{pa(j)}, e) = F^{-1}_{Z_j \mid Z_{pa(j)} = \tilde{z}_{pa(j)}}(e)$$

$$\varepsilon_1, \dots, \varepsilon_p \sim U[0,1]$$

$F_{Z_j \mid Z_{pa(j)} = \tilde{z}_{pa(j)}}$ is the distribution fn of $Z_j \mid Z_{pa(j)} = \tilde{z}_{pa(j)}$.

But P will be Markov w.r.t many DAGs (e.g. any DAG whose skeleton is the complete graph).
So e.g. $Z_1 \perp\!\!\!\perp Z_2$ can be expressed

$$Z_1 = 0 \times Z_2 + \varepsilon_1, \qquad Z_2 = \varepsilon_2 \qquad Z_1 \leftarrow Z_2$$

Defn P satisfies causal minimality w.r.t. DAG $\mathcal{G}$ if it is Markov w.r.t. $\mathcal{G}$ but not Markov w.r.t. any proper subgraph of $\mathcal{G}$.

Markov equivalent DAGs

Two different DAGs may satisfy the same collection of d-separations, e.g. $Z_1 \leftarrow Z_2$ & $Z_1 \rightarrow Z_2$. For DAG $\mathcal{G}$,
let $\mathcal{M}(\mathcal{G}) = \{$distributions P : P is Markov w.r.t. $\mathcal{G}\}$.

Defn Two DAGs are Markov equivalent if $\mathcal{M}(\mathcal{G}_1) = \mathcal{M}(\mathcal{G}_2)$.