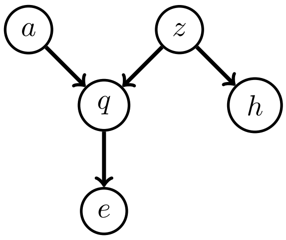
The purpose of this additional sheet is to provide more practice and exam preparation material. N.B. The tutors are not required to work through this material in the tutorial.

Exercise 1. ***More on ordered and local Markov properties, d-separation***

We continue with the investigation of the graph below



(a) Why can the ordered or local Markov property not be used to check whether $a \perp\!\!\!\perp h \mid e$ may hold?

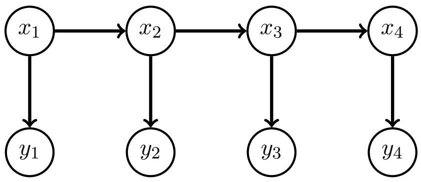
(b) Use d-separation to check whether $a \perp\!\!\!\perp h \mid e$ holds.

(c) The independency relations obtained via the ordered and local Markov property include $a \perp\!\!\!\perp \{z, h\}$. Verify the independency using d-separation.

(d) Determine the Markov blanket of z.

Exercise 2. ***Hidden Markov models***

This exercise is about directed graphical models that are specified by the following DAG:



These models are called "hidden" Markov models because we typically assume to only observe the y_i and not the x_i that follow a Markov model.

(a) Show that all probabilistic models specified by the DAG factorise as

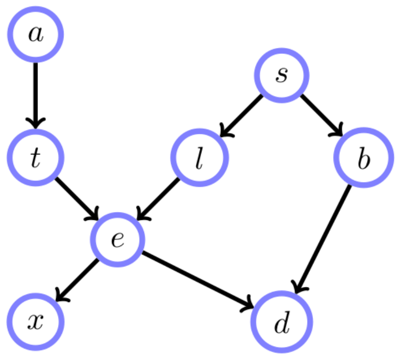
$$p(x_1, y_1, x_2, y_2, \dots, x_4, y_4) = p(x_1)p(y_1|x_1)p(x_2|x_1)p(y_2|x_2)p(x_3|x_2)p(y_3|x_3)p(x_4|x_3)p(y_4|x_4)$$

(b) Derive the independencies implied by the ordered Markov property with the topological ordering $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$

(c) Derive the independencies implied by the ordered Markov property with the topological ordering $(x_1, x_2, \dots, x_4, y_1, \dots, y_4)$.

Exercise 3. ***More on the chest clinic (based on Barber's exercise 3.3)***

The directed graphical model in Figure 1 is the "Asia" example of Lauritzen and Spiegelhalter (1988). It concerns the diagnosis of lung disease (T=tuberculosis or L=lung cancer). In this model, a visit to some place in A=Asia is thought to increase the probability of tuberculosis.



$x =$ Positive X-ray
$d =$ Dyspnea (Shortness of breath)
$e =$ Either Tuberculosis or Lung Cancer
$t =$ Tuberculosis
$l =$ Lung Cancer
$b =$ Bronchitis
$a =$ Visited Asia
$s =$ Smoker

Figure 1: Graphical model for Exercise 3 (Barber Figure 3.15).

(a) Explain which of the following independence relationships hold for all distributions that factorise over the graph.

1. $a \perp\!\!\!\perp s \mid l$
2. $a \perp\!\!\!\perp s \mid l, d$

(b) Let g be a (deterministic) function of x and t. Is the expected value $\mathbb{E}[g(x,t) \mid l, b]$ equal to $\mathbb{E}[g(x,t) \mid l]$?

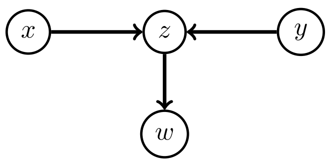
Exercise 4. ***Independencies***

This exercise is on further properties and characterisations of statistical independence.

(a) Without using d-separation, show that $x \perp\!\!\!\perp \{y, w\} \mid z$ implies that $x \perp\!\!\!\perp y \mid z$ and $x \perp\!\!\!\perp w \mid z$. *Hint: use the definition of statistical independence in terms of the factorisation of pmfs/pdfs.*

(b) For the directed graphical model below, show that the following two statements hold without using d-separation:

$$x \perp\!\!\!\perp y \quad \text{and} \tag{1}$$
$$x \not\perp\!\!\!\perp y \mid w \tag{2}$$



The exercise shows that not only conditioning on a collider node but also on one of its descendents activates the trail between x and y. You can use the result that $x \perp\!\!\!\perp y | w \Leftrightarrow p(x,y,w) = a(x,w)b(y,w)$ for some non-negative functions $a(x,w)$ and $b(y,w)$.