The purpose of the tutorials is twofold: First, they help you better understand the lecture material. Secondly, they provide exam preparation material. You are not expected to complete all questions before the tutorial sessions. Start early and do as many as you have time for.

Exercise 1. Visualising and analysing Gibbs distributions via undirected graphs

We here consider the Gibbs distribution

\[ p(x_1, \ldots, x_5) \propto \phi_{12}(x_1, x_2)\phi_{13}(x_1, x_3)\phi_{14}(x_1, x_4)\phi_{23}(x_2, x_3)\phi_{25}(x_2, x_5)\phi_{45}(x_4, x_5) \]

(a) Visualise it as an undirected graph.

(b) What are the neighbours of \( x_3 \) in the graph?

(c) Do we have \( x_3 \perp \perp x_4 \mid x_1, x_2 \)?

(d) What is the Markov blanket of \( x_4 \)?

(e) On which minimal set of variables \( A \) do we need to condition to have \( x_1 \perp \perp x_5 \mid A \)?

Exercise 2. Factorisation and independencies for undirected graphical models

We here consider the graph in Figure 1.

\[ \]

(a) What is the set of Gibbs distributions that are induced by the graph?

(b) Let \( p \) be a pdf that factorises according to the graph. Can we expect that \( p(x_3 \mid x_2, x_4) = p(x_3 \mid x_4) \)?

(c) Explain why \( x_2 \perp \perp x_5 \mid x_1, x_3, x_4, x_6 \) holds.

(d) Assume you would like to approximate \( E(x_1 x_2 x_5 \mid x_3, x_4) \), i.e. the expected value of the product of \( x_1, x_2, \) and \( x_5 \) given \( x_3 \) and \( x_4 \), with a sample average. Do you need to have joint observations for all five variables, i.e. of the tuples \( (x_1, x_2, x_3, x_4, x_5) \)?

Figure 1: Graph for Exercise 2
Exercise 3. *Undirected graphical model with pairwise potentials*

We here consider Gibbs distributions where the factors only depend on two variables at a time. The probability density or mass functions over \(d\) random variables \(x_1, \ldots, x_d\) then take the form

\[
p(x_1, \ldots, x_d) \propto \prod_{i \leq j} \phi_{ij}(x_i, x_j)
\]

These models are typically called pairwise Markov networks.

(a) Let \(p(x_1, \ldots, x_d) \propto \exp\left(-\frac{1}{2}x^\top A x - b^\top x\right)\) where \(A\) is symmetric and \(x = (x_1, \ldots, x_d)^\top\). What are the corresponding factors \(\phi_{ij}\) for \(i \leq j\)?

(b) For \(p(x_1, \ldots, x_d) \propto \exp\left(-\frac{1}{2}x^\top A x - b^\top x\right)\), show that \(x_i \perp \perp x_j \mid \{x_1, \ldots, x_d\} \setminus \{x_i, x_j\}\) if the \((i, j)\)-th element of \(A\) is zero.

Exercise 4. *Restricted Boltzmann machine (based on Barber Exercise 4.4)*

The restricted Boltzmann machine is an undirected graphical model for binary variables \(v = (v_1, \ldots, v_n)^\top\) and \(h = (h_1, \ldots, h_m)^\top\) with a probability mass function equal to

\[
p(v, h) \propto \exp\left(v^\top W h + a^\top v + b^\top h\right),
\]

where \(W\) is a \(n \times m\) matrix. Both the \(v_i\) and \(h_i\) take values in \(\{0, 1\}\). The \(v_i\) are called the “visibles” variables since they are assumed to be observed while the \(h_i\) are the hidden variables since it is assumed that we cannot measure them.

(a) Use graph separation to show that the joint conditional \(p(h \mid v)\) factorises as

\[
p(h \mid v) = \prod_{i=1}^m p(h_i \mid v).
\]

(b) Show that

\[
p(h_i = 1 \mid v) = \frac{1}{1 + \exp\left(-b_i - \sum_j W_{ji} v_j\right)}
\]

where \(W_{ji}\) is the \((ji)\)-th element of \(W\), so that \(\sum_j W_{ji} v_j\) is the inner product between the \(i\)-th column of \(W\) and \(v\).

(c) Use a symmetry argument to show that

\[
p(v \mid h) = \prod_i p(v_i \mid h) \quad \text{and} \quad p(v_i = 1 \mid h) = \frac{1}{1 + \exp\left(-a_i - \sum_j W_{ij} h_j\right)}
\]