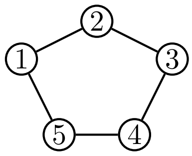
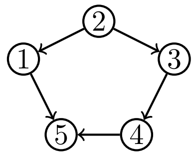
Practice Problems

26-30 May 2025

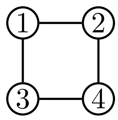
Event: Finnish Summer School on Probability and Statistics
Author: Liam Solus (solus@kth.se)

1. (*Markov properties.*)Write down each of the Markov properties for the following graphs:

2. (*Different graph types yield different models.*).

   (a) Show that the collection of all positive distributions on (X_1, X_2, X_3, X_4) that are continuous on $\mathcal{X}$ and satisfy $X_1 \perp\!\!\!\perp X_4 \mid X_{2,3}$ and $X_2 \perp\!\!\!\perp X_3 \mid X_{1,4}$ is the undirected Gaussian graphical model $\mathcal{M}(G)$ where G is the graph



   (b) Show that there is no DAG G on node set $[4] = \{1, 2, 3, 4\}$ such that $\mathcal{M}(G)$ is the set of all distributions satisfying $X_1 \perp\!\!\!\perp X_4 \mid X_{2,3}$ and $X_2 \perp\!\!\!\perp X_3 \mid X_{1,4}$.

3. (*The conditional independence axioms.*) Suppose that P is a distribution on $X = (X_i)_{i\in[m]}$ with probability density (or mass) function $f_X(x) > 0$ for all $x = (x_1, \dots, x_m) \in \mathcal{X}$, and let $A, B, C, D \subseteq [m]$. Prove the following:

   (1) (Symmetry) If P satisfies $X_A \perp\!\!\!\perp X_B \mid X_C$ then P satisfies $X_B \perp\!\!\!\perp X_A \mid X_C$.

   (2) (Decomposition) If P satisfies $X_A \perp\!\!\!\perp X_{B\cup D} \mid X_C$ then P satisfies $X_A \perp\!\!\!\perp X_B \mid X_C$.

   (3) (Weak Union) If P satisfies $X_A \perp\!\!\!\perp X_{B\cup D} \mid X_C$ then P satisfies $X_A \perp\!\!\!\perp X_B \mid X_{C\cup D}$.

   (4) (Contraction) If P satisfies $X_A \perp\!\!\!\perp X_B \mid X_{C\cup D}$ and $X_A \perp\!\!\!\perp X_D \mid X_C$ then P satisfies $X_A \perp\!\!\!\perp X_{B\cup D} \mid X_C$.

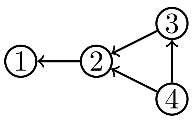
   (5) (Intersection) If P satisfies $X_A \perp\!\!\!\perp X_B \mid X_{C\cup D}$ and $X_A \perp\!\!\!\perp X_C \mid X_{B\cup D}$ then $X_A \perp\!\!\!\perp X_{B\cup C} \mid X_D$.

4. (*Conditional independence in Gaussian distributions.*) Let $G = ([m], E)$ be an undirected graph, and let $X = (X_i)_{i\in[m]} \sim \mathrm{N}(0, \Sigma)$. Show that $P \in \mathcal{M}(G)$ if and only if $\mathrm{Cov}[X_i, X_j \mid X_{[m]\setminus\{i,j\}}] = 0$ for all $i - j \notin E$.

5. *(The relationship between directed pairwise and local Markov properties.)*

   (a) Suppose that a distribution P satisfies the directed local Markov property for a DAG G. Show that P satisfies the directed pairwise Markov property for G.

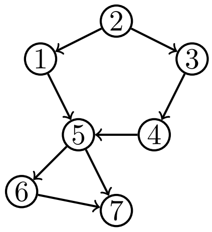
   (b) Find a distribution on four binary random variables that satisfies the directed pairwise Markov property with respect to the graph



   but does not satisfy the directed local Markov property. Show that this distribution does not factorize according to the given graph.

6. (*Markov equivalence.*)

   (a) What graphs are in the Markov equivalence class of the following graph?



   (b) Let $I_m = ([m+1], E_m)$ be the graph with edge set

$$E_m = \{1 - 2, 2 - 3, \ldots, m - (m+1)\}$$

   How many Markov equivalence classes are there on the directed acyclic graphs having skeleton I_m.

7. (*Learning a Markov equivalence class via CI tests.*) Let $G = ([m], E)$ be a DAG.

   (a) Let $i, j \in [m]$. Show that i and j are connected by an edge in E if and only if there is no set $C \subseteq [m] \setminus \{i, j\}$ such that i and j are d-separated given C in G.

   (b) Let $i, j, k \in [m]$, and suppose that i, k and j, k are each connected by an edge in G while i, j are not. Show that $i \to k \leftarrow j$ is in G if and only if there exists a set $C \subseteq [m] \setminus \{i, j\}$ not containing k such that i and j are d-separated given C in G.

   (c) Let $X^{(1)}, \ldots, X^{(n)}$ be a random sample from some distribution $P \in \mathcal{M}(G)$ for some unknown DAG G. Use parts (a) and (b) to implement an algorithm for estimating G from this random sample using statistical hypothesis tests for conditional independence. Explain why your algorithm works for different data types (e.g., Gaussian, discrete, etc.).

8. (*Causal graph discovery under equal error variance assumption.*) Let $G = (|m], E)$ be a DAG and let $\mathcal{M}_{\mathrm{ev}}(G)$ denote the set of all distributions over $X = (X_i)_{i \in [m]}$ defined by the system of equations

$$X_i = \sum_{j \in \mathrm{pa}_G(i)} \lambda_{ij} X_j + \varepsilon_i \qquad \forall i \in [m],$$

   for some $\lambda_{ij} \in \mathbb{R}$, some $\omega \in \mathbb{R}_{>0}$ and some mutually independent $\varepsilon_1, \ldots, \varepsilon_m$ where $\varepsilon_i \sim \mathrm{N}(0, \omega)$ for all $i \in [m]$. Suppose $X^{(1)}, \ldots, X^{(n)}$ is a random sample from some distribution $P \in \mathcal{M}_{\mathrm{ev}}(G)$ where G is unknown. Implement an algorithm that learns G from this random sample. Explain why your algorithm learns a graph G as opposed to an equivalence class of graphs.

9. (*Using computational algebra to learn model constraints.*) Let M denote the collection of all Gaussian distributions $\mathrm{N}(0, \Sigma)$ on $X = (X_1, X_2, X_3)$ where

$$\Sigma^{-1} \in \mathcal{M} = \{K = [k_{i,j}]_{i,j=1}^{3} \in \mathrm{PD}_3 : k_{1,3} = 0, k_{1,1} = k_{3,3}\}.$$

   (a) Use `Macaulay2` (online at https://www.unimelb-macaulay2.cloud.edu.au/#home) to find a basis for the vanishing ideal $I(\mathcal{M})$.

   (b) Give the elements of the basis found in part (a) a statistical interpretation.

   (c) Give hypothesis tests for deciding if a distribution P belongs to $\mathcal{M}$ based on a random sample.