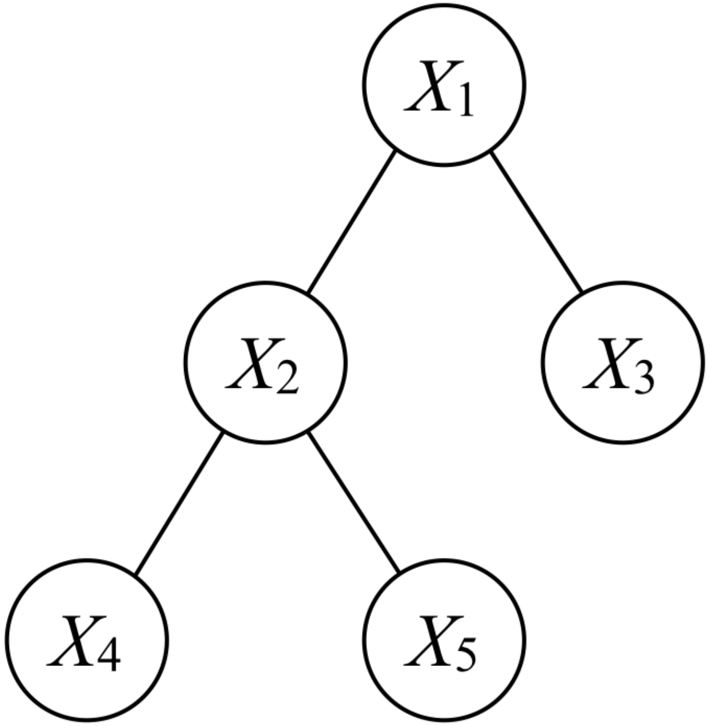
Exercise: Graphical Models

Exercise: Graphical Models

5/5 points (graded)

Consider an undirected graphical model with the following graph:



- Which of the following probability distributions can the above graphical model encode? (Select all that are possible from the list below.)

- [x] A probability distribution in which X_1, X_2, X_3, X_4, X_5 are independent: $p_{X_1,X_2,X_3,X_4,X_5} = p_{X_1} p_{X_2} p_{X_3} p_{X_4} p_{X_5}$

- [x] A probability distribution with factorization $p_{X_1,X_2,X_3,X_4,X_5} = p_{X_1} p_{X_2|X_1} p_{X_3|X_1} p_{X_4|X_2} p_{X_5|X_2}$

- [x] A probability distribution with factorization $p_{X_1,X_2,X_3,X_4,X_5} = p_{X_1} p_{X_2} p_{X_3|X_1} p_{X_4|X_2} p_{X_5}$ ✔
- [ ] A probability distribution with factorization $p_{X_1,X_2,X_3,X_4,X_5} = p_{X_3} p_{X_1|X_4} p_{X_4} p_{X_2|X_1} p_{X_5}$

✔

Suppose there are n random variables $X_1, \ldots, X_n$ and each random variable X_i takes on k different values.

- If there's no known structure, how many numbers are needed to specify the full joint distribution of $X_1, \ldots, X_n$?

- ( ) $\mathcal{O}(n^k)$
- (•) $\mathcal{O}(k^n)$ ✔
- ( ) $\mathcal{O}(k \cdot n)$
- ( ) $\mathcal{O}(k + n)$

- Suppose now that we know the joint distribution of the n random variables actually corresponds to an undirected graphical model with m edges (no loops of course). How many numbers are needed to store all the potential tables? (Pick the choice that is the smallest in big O.)

  - ○ $\mathcal{O}(k^n)$
  - ○ $\mathcal{O}(k \cdot n \cdot m)$
  - ○ $\mathcal{O}(n \cdot k^2 + m \cdot k^2)$
  - ◉ $\mathcal{O}(n \cdot k + m \cdot k^2)$ ✔

- When there are no loops in a graph with n nodes, what is the largest that the number of edges m can be in terms of n?

  In this part, please provide your answer as a mathematical formula (and not as Python code). Use ^ for exponentiation, e.g., x^2 denotes x^2. Explicitly include multiplication using *, e.g. x*y is xy.

  n-1 ✔ **Answer:** n-1

  $n - 1$

Let's take a look a closer look at the normalization constant Z.

- Let's return to our first example where we have two random variables X_1 and X_2 that are independent. We represent it as a graphical model

$$p_{X_1,X_2} = \frac{1}{Z}\phi_1(x_1)\phi_2(x_2),$$

where now we set $\phi_1(x_1) = p_{X_1}(x_1)$ and $\phi_2(x_2) = 2p_{X_2}(x_2)$.

What is Z in this case?

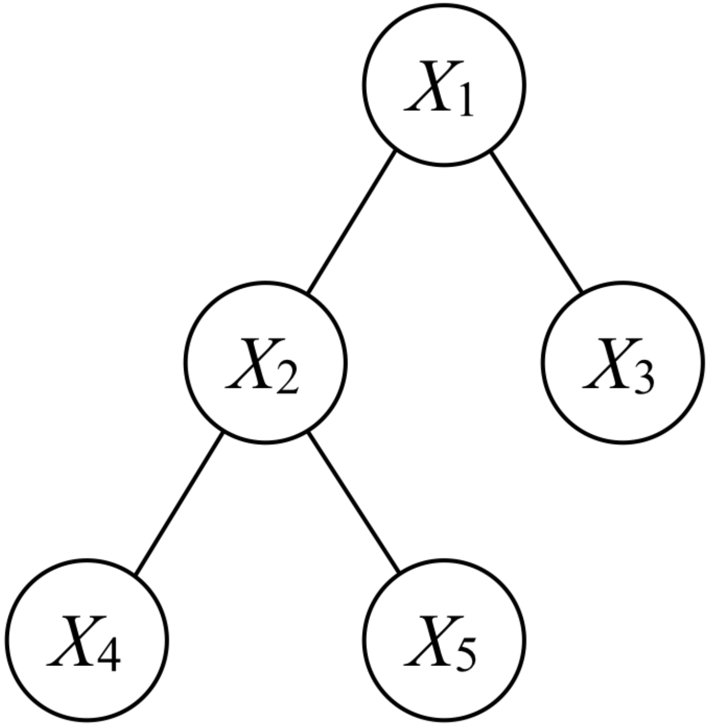
2

Answer: 2

Very importantly: this problem tells us that there are actually infinitely many ways we can specify the potential tables because we can always scale all the entries by the same positive constant! What changes is Z.

Solution:

Consider an undirected graphical model with the following graph:



- Which of the following probability distributions can the above graphical model encode? (Select all that are possible from the list below.)

  Solution: The graphical model can encode a distribution when $X_1, \ldots, X_n$ are independent: just set all the $\psi_{i,j}$'s to output 1 for all inputs, which effectively means that the edges need not be there.

  A probability distribution with factorization $p_{X_1,X_2,X_3,X_4,X_5} = p_{X_1} p_{X_2|X_1} p_{X_3|X_1} p_{X_4|X_2} p_{X_5|X_2}$ can indeed be encoded; by just looking at the pairwise factors we see that they match up with the edges in the graph.

A probability distribution with factorization $p_{X_1,X_2,X_3,X_4,X_5} = p_{X_1} p_{X_2} p_{X_3|X_1} p_{X_4|X_2} p_{X_5}$ can also be encoded; the pairwise factors are a subset of what edges are present in the graph.

A probability distribution with factorization $p_{X_1,X_2,X_3,X_4,X_5} = p_{X_3} p_{X_1|X_4} p_{X_4} p_{X_2|X_1} p_{X_5}$ can *not* be encoded as there is a pairwise factor that depends on both X_1 and X_4, and there is no edge $(1,4)$ in the graph.

Suppose there are n random variables $X_1, \ldots, X_n$ and each random variable X_i takes on k different values.

- If there's no known structure, how many numbers are needed to specify the full joint distribution of $X_1, \ldots, X_n$?

  Solution: There are n variables, each of which takes of k possible values, so the number of entries in the joint probability table is

$$\underbrace{k \cdot k \cdots k}_{n \text{ times}} = k^n = \boxed{\mathcal{O}(k^n)}.$$

- Suppose now that we know the joint distribution of the n random variables actually corresponds to an undirected graphical model with m edges (no loops of course). How many numbers are needed to store all the potential tables? (Pick the choice that is the smallest in big O.)

  Solution: There are n nodes each with a node potential table that can at most take $\mathcal{O}(k)$ space (we could possibly treat node potentials that are all 1's differently but here it suffices to get a worst-case upper bound), and m edges each with a pairwise potential table that takes $\mathcal{O}(k^2)$ space. So in total we need $\mathcal{O}(n \cdot k + m \cdot k^2)$ space.

- When there are no loops in a graph with n nodes, what is the largest that the number of edges m can be in terms of n?

  Solution: For a graph with n nodes, and no loops, trees have the most number of edges: $m = \boxed{n-1}$ edges. Any additional edge will cause a loop to form.

Let's take a look a closer look at the normalization constant Z.

- Let's return to our first example where we have two random variables X_1 and X_2 that are independent. We represent it as a graphical model

$$p_{X_1,X_2} = \frac{1}{Z}\phi_1(x_1)\phi_2(x_2),$$

  where now we set $\phi_1(x_1) = p_{X_1}(x_1)$ and $\phi_2(x_2) = 2p_{X_2}(x_2)$.

  What is Z in this case?

  Solution: If we're scaling all the entries of p_{X_2} by 2, then Z, which was 1 before and ensured that the distribution sums to 1, now becomes $Z = \boxed{2}$ to counteract p_{X_2} being scaled by 2.

  Very importantly: this problem tells us that there are actually infinitely many ways we can specify the potential tables because we can always scale all the entries by the same positive constant! What changes is Z.

Submit You have used 1 of 5 attempts

✔ Correct (5/5 points)