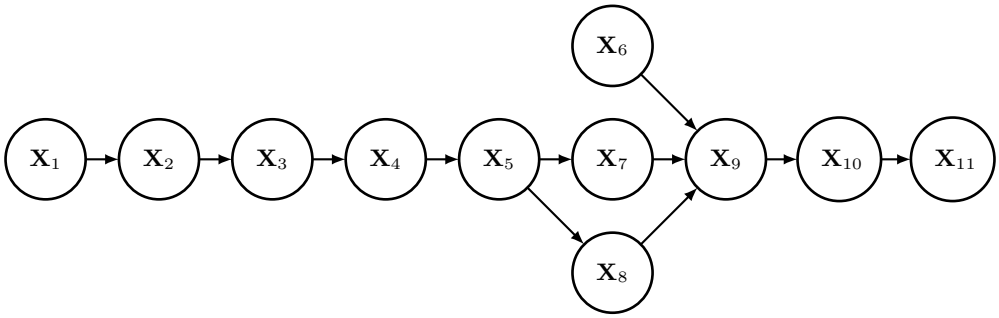
SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

50.007 Machine Learning, Fall 2021
Homework 5

Due Thursday 9 December 2021, 5pm

This homework will be graded by Zhang Qi

In this homework, we would like to look at the Bayesian Networks. You are given a Bayesian network as below. All nodes can take 2 different values: $\{1, 2\}$.



Question 1. Without knowing the actual value of any node, are node $\mathbf{X}_1$ and $\mathbf{X}_6$ independent of each other? What if we know the value of node $\mathbf{X}_5$ and $\mathbf{X}_{10}$? *(5 points)*

Question 2. What is the *effective* number of parameters needed to for this Bayesian network? What would be the *effective* number of parameters for the same network if node $\mathbf{X}_3$, $\mathbf{X}_8$ and $\mathbf{X}_9$ can take 5 different values: $\{1, 2, 3, 4, 5\}$, and all other nodes can only take 4 different values: $\{1, 2, 3, 4\}$? *(5 points)*

Question 3. If we have the following probability tables for the nodes. Compute the following probabilities. Clearly write down all the necessary steps.

(a) Calculate the following conditional probability:

$$P(\mathbf{X}_3 = 1 | \mathbf{X}_4 = 2)$$

(6 points)

(b) Calculate the following conditional probability:

$$P(\mathbf{X}_5 = 2 | \mathbf{X}_2 = 1, \mathbf{X}_{11} = 2, \mathbf{X}_1 = 1)$$

(9 points)

(*Hint: find a short answer. The values in some of the probability tables may reveal some useful information.*)

	$\mathbf{X}_1$	
	1	2
	0.5	0.5

$\mathbf{X}_1$	$\mathbf{X}_2$	
	1	2
1	0.2	0.8
2	0.3	0.7

$\mathbf{X}_2$	$\mathbf{X}_3$	
	1	2
1	0.3	0.7
2	0.3	0.7

$\mathbf{X}_3$	$\mathbf{X}_4$	
	1	2
1	0.1	0.9
2	0.5	0.5

$\mathbf{X}_4$	$\mathbf{X}_5$	
	1	2
1	0.5	0.5
2	0.6	0.4

	$\mathbf{X}_6$	
	1	2
	0.6	0.4

$\mathbf{X}_5$	$\mathbf{X}_7$	
	1	2
1	0.2	0.8
2	0.3	0.7

$\mathbf{X}_5$	$\mathbf{X}_8$	
	1	2
1	0.8	0.2
2	0.7	0.3

$\mathbf{X}_6$	$\mathbf{X}_7$	$\mathbf{X}_8$	$\mathbf{X}_9$	
			1	2
1	1	1	0.8	0.2
1	1	2	0.1	0.9
1	2	1	0.9	0.1
1	2	2	0.7	0.3
2	1	1	0.3	0.7
2	1	2	0.2	0.8
2	2	1	0.2	0.8
2	2	2	0.9	0.1

$\mathbf{X}_9$	$\mathbf{X}_{10}$	
	1	2
1	0.8	0.2
2	0.8	0.2

$\mathbf{X}_{10}$	$\mathbf{X}_{11}$	
	1	2
1	0.7	0.3
2	0.8	0.2

Question 4.

(a) Now, assume we do not have any knowledge about the probability tables for the nodes in the network, but we have the following 12 observations/samples. Find a way to estimate the probability tables associated with the nodes $\mathbf{X}_7$ and $\mathbf{X}_9$ respectively. *(6 points)*

$\mathbf{X}_1$	$\mathbf{X}_2$	$\mathbf{X}_3$	$\mathbf{X}_4$	$\mathbf{X}_5$	$\mathbf{X}_6$	$\mathbf{X}_7$	$\mathbf{X}_8$	$\mathbf{X}_9$	$\mathbf{X}_{10}$	$\mathbf{X}_{11}$
1	1	2	2	2	1	1	1	2	1	1
1	2	1	1	2	1	1	1	1	1	2
2	2	2	1	2	2	1	1	1	2	1
1	1	2	1	2	1	1	2	1	2	2
1	2	1	1	1	1	2	2	2	1	1
2	2	1	2	1	2	2	1	1	1	2
2	1	2	2	1	2	1	2	2	2	1
2	2	2	1	2	1	2	2	1	2	2
1	1	1	1	2	2	1	1	1	1	1
1	1	1	1	2	1	1	1	2	1	2
1	2	1	2	2	1	2	1	1	1	2
2	2	1	2	1	2	2	2	2	1	1

(b) Based on the above observations, you would like to find a good Bayesian network structure to model the data. You started with the initial structure shown on the previous page, and decided to delete the edge between $\mathbf{X}_{10}$ and $\mathbf{X}_{11}$. Is the resulting new structure (after deleting the single edge between $\mathbf{X}_{10}$ and $\mathbf{X}_{11}$ from the original graph) better than the original structure in terms of BIC score? Clearly explain the reason. *(9 points)*

(Hint: Try to find a short answer.)