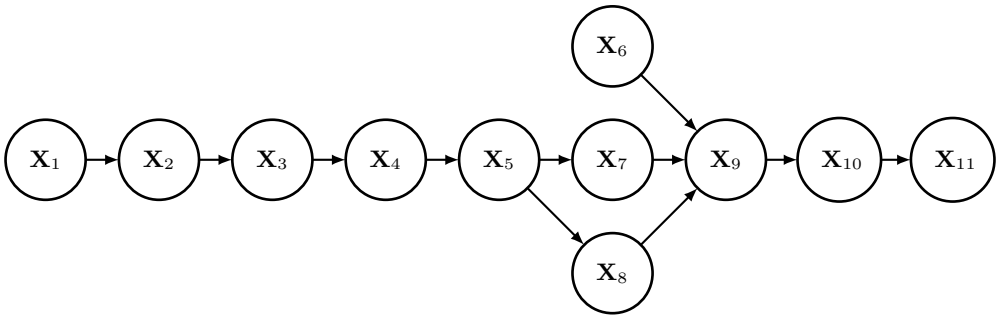
SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

01.112 Machine Learning, Fall 2018
Homework 5

Due Friday 7 Dec 2018, 5pm

This homework will be graded by Ngo Van Mao

Question 1 Consider the Bayesian network below, where we have 11 variables.



1. (4 points) Assume all variables are taking values from $\{1, 2, 3\}$. What is the number of *free parameters* for? What if we assume all variables are taking values from $\{1, 2, 3, 4\}$? Show detailed steps that lead to your answers.

2. (4 points) What is the Markov blanket for the variable X_1 in the Bayesian network? What is the Markov blanket for the variable X_7?

3. (6 points) Are X_1 and X_6 independent or dependent of each other if no other variable is given? Why? Are X_1 and X_6 independent or dependent of each other if both X_7 and X_{10} are given? Why?

4. (8 points) Now, assume the probability tables for all nodes are shown below (next page):

   Calculate the following conditional probability:

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

   (*Hint: find a short answer.*)

5. (8 points) Calculate the following conditional probability based on the same probability tables.

   $$P(\mathbf{X}_5 = 2|\mathbf{X}_3 = 1, \mathbf{X}_{11} = 2, \mathbf{X}_1 = 1)$$

   (*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

$\mathbf{X}_6$	$\mathbf{X}_7$	$\mathbf{X}_8$	$\mathbf{X}_9$	
			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

$\mathbf{X}_5$	$\mathbf{X}_7$	
	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

$\mathbf{X}_9$	$\mathbf{X}_{10}$	
	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

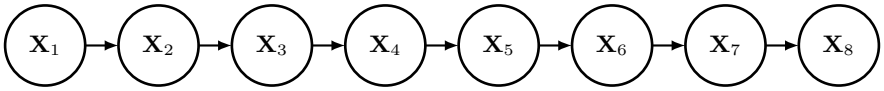


Figure 1: G_1

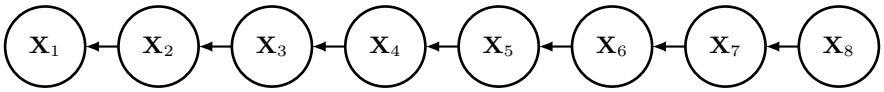


Figure 2: G_2

Question 2 (10 points) Now consider the above two Bayesian network structures, where all variables are binary. In other words, they are taking values from $\{1, 2\}$.

Now, you would like to use BIC as the criterion for selecting a better structure of the Bayesian network between G_1 and G_2 based on a large collection of samples. Construct a case (i.e., provide a collection of samples, where each sample is of the form "$X_1 = 1$, $X_2 = 2$,...,$X_8 = 2$", for example) where the final BIC of the first structure G_1 would be strictly higher than G_2. If you believe no such case exists, clearly explain why.