**Birla Institute of Technology & Science, Pilani**

**Work Integrated Learning Programmes Division**

**Second Semester 2021-2022**

**Comprehensive Examination**

**(EC-3 Regular)**

Course No. : DSECLZG526

Course Title : Probabilistic Graphical Models

Nature of Exam : Open Book

Weightage : 40% (As per Course Handout)

No. of Pages = 3

# No. of Questions = 6

Duration : 2.5 Hours

Date of Exam : 01/10/2022 (AN)

Note to Students:

1. Please follow all the *Instructions to Candidates* given on the cover page of the answer book.
2. All parts of a question should be answered consecutively. Each answer should start from a fresh page.
3. Assumptions made if any, should be stated clearly at the beginning of your answer.

Q.1 Use the Chow-Liu algorithm to obtain the best tree-approximation to the distribution represented by the following empirical data:

|  |  |  |  |
| --- | --- | --- | --- |
| A | B | C | D |
| 1 | 1 | 0 | 1 |
| 1 | 1 | 0 | 0 |
| 1 | 0 | 1 | 1 |
| 0 | 1 | 1 | 0 |
| 1 | 1 | 0 | 0 |

[10 Marks]

Answer:

We compute the mutual information between every pair of variables which become weights of edges in a graph. We then find the maximum weight spanning tree in the complete graph G consisting of pairwise weights.

The mutual information can be computed as follows for :

.

Similarly

.

.

.

.

.

The graph we get is as below:



The final maximum weight spanning tree is as below:



Suggested Marking Scheme:

Calculation of all Mutual Information numbers 🡪 7 Marks

Finding maximum weight spanning tree 🡪 3 Marks

Q.2 For the Markov chain below, find the stationary distribution. The transition probabilities are as follows:

A

B

C

[6 Marks]

Answer:

The transition probability matrix can be written as follows To find the stationary distribution we seek a non-zero vector such that and We see that such a vector is .

Suggested Marking Scheme: Setting up the probability transition matrix 🡪 2 Marks

Solving for 🡪 4 marks

Q.3 Consider the distribution on the binary variables such that evaluates to for each of the combinations in the table below and zero for all other combinations. Determine if factors according to the graph below (on the right). Give a mathematical justification for your answer.

4

3

2

1

|  |  |
| --- | --- |
| 0000 | 1/8 |
| 1000 | 1/8 |
| 1100 | 1/8 |
| 1110 | 1/8 |
| 0001 | 1/8 |
| 0011 | 1/8 |
| 0111 | 1/8 |
| 1111 | 1/8 |

[6 Marks]

Answer: Let us assume that factorizes according to the given graph . We have .

Now and implies that or or both are zero.

contradicts with and contradicts with , so we end up with a contradiction either way. Thus cannot be factorized according to the given graph

Suggested Marking Scheme:

Idea of contradiction proof assuming that factorizes according to 🡪 2 Marks

Complete proof using contradiction 🡪 4 Marks

Q.4 Is it possible for a Bayesian network on binary-valued variables to have edges such that its minimal undirected I-map will have edges? If it is possible, draw such a Bayesian network and the equivalent Markov network for and write the joint distribution for the Bayesian network in factorized form on and determine the number of nonredundant parameters needed to characterize the factorized joint distribution in terms of . Otherwise, explain why it is not possible. [6 Marks]

Answer:

Yes, it is possible for a Bayesian network on binary-valued variables to have edges such that its minimal undirected I-map will have edges. The Bayesian network should have one of the nodes as a child of the remaining nodes. For example the network could have as a child of The minimal undirected I-map for this network will be a moralized graph where we will marry the parents of and add edges between every pair of parents.



The joint distribution of the Bayesian network in factorized form can be written as In terms of nonredundant parameters, we need one each for each of to specify the probability and in order to specify , we need a parameter for the probability of given every possible combination of . There are combinations for these variables, so in total the number of nonredundant parameters are

Suggested Marking Scheme:

Drawing minimal undirected I-map 🡪 3 Marks

Calculating nonredundant parameters 🡪 3 Marks

Q.5 Apply the variable-elimination algorithm on the graph below using the elimination orderings and . List all the intermediate steps in the computation. Assume that each of the variables is binary-valued, and find out the size of the largest factor formed in each case. [6 Marks]

B

A

E

D

C

H

G

F

Answer The Bayesian factorization for the joint probability distribution of the given network can be written as

Let us consider the elimination order first.

Eliminating : and .

Eliminating :

Eliminating :

Eliminating : ,

Eliminating :

Eliminating :

Eliminating :

Eliminating :

The largest factor formed is the one on the variables which has rows in it.

Let us now consider the elimination ordering .

Eliminating : and .

Eliminating :

Eliminating :

Eliminating :

Eliminating :

Eliminating :

Eliminating :

Eliminating :

The largest factor formed in this case is on 3 variables, which means that the largest factor has rows.

Suggested Marking Scheme

First elimination ordering 🡪 3 Marks

Second elimination ordering 🡪 3 Marks

Q.6 Prior information about the parameters of a biased coin with sample space suggests that the parameter vector obeys a Dirichlet distribution with hyperparameters 2 and 3 respectively. Let stand for the outcome and for the outcome . What is What is the smallest number of samples that we need in order for us to be able to conclude that where ?

[6 Marks]

Answer

We have

Further , where . We need which means that or .

The smallest value for is 0 which means . Thus is the smallest number of samples we need in order to conclude that .

Suggested Marking Scheme

Getting 🡪 2 Marks

Getting the smallest number of samples right 🡪 4 Marks