Lista 03 de Exercícios - PGM

Renato Assunção - DCC, UFMG

Setembro de 2017

The next questions refer to the student BN reproduced in Figure 1. Remember that we are assuming that the joint distribution of the random vector $\mathbf{X} = (D, I, G, S, L)$ factorizes over the graph. That is, we are assuming that, for any instantiation (d, i, g, s, l) of the random vector $\mathbf{X}$, we have

$$\mathbb{P}(D = d, I = i, G = g, S = s, L = l) = \mathbb{P}(D = d)\mathbb{P}(I = i)\mathbb{P}(S = s|I = i)\mathbb{P}(G = g|I = i, D = d)\mathbb{P}(L = l|G = g) \tag{1}$$

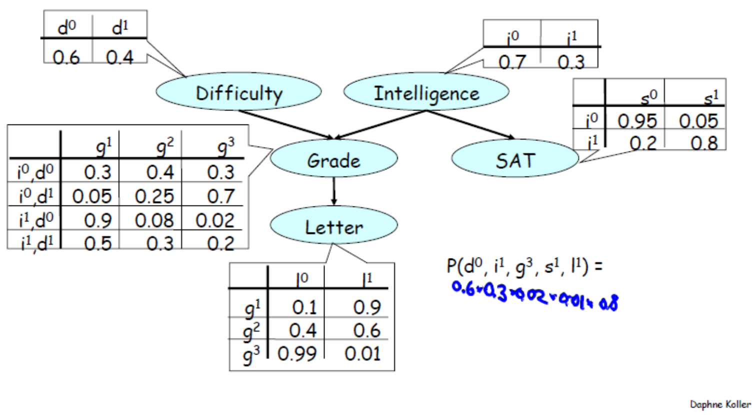

Figura 1: Student BN with the associated CPTs for each node.

1. Show numerically that (1) is a valid probability distribution. That is, argue convincingly that it assigns values greater or equal to zero to each one of the $2^4 \times 3 = 48$ instantiations of $\mathbf{X}$ and that it is equal to 1 when summing over all possible values. If needed, watch the Daphne Koller video *Semantics & Factorization* at 9:58 in the platform *coursera* or at `https://www.youtube.com/watch?v=6ODl1rxoT14`.

2. Using the same "pushing in" summation sign trick from the previous exercise, show that the marginal distribution of the random variable D is given by the distribution given to the node D in the Student BN. That is, show that, by summing out the joint distribution (1) over all the possible values of the remaining variables in the BN (I, G, S, L), we have $\mathbb{P}(D = d^0) = 0.6$ and $\mathbb{P}(D = d^1) = 0.4$.

3. Obtain the marginal distribution of S by summing out the joint distribution (1) over all other variables. You should find $\mathbb{P}(S = s^0) = 0.725$ and $\mathbb{P}(S = s^1) = 1 - 0.725$.

4. Show numerically that $D \perp I$ by showing that $\mathbb{P}(D = d, I = i) = \mathbb{P}(D = d)\mathbb{P}(I = i)$ for all values of d and i. In particular, what is the value of $\mathbb{P}(D = d^0, I = i^1)$?

5. Effect of v-structure: Show numerically that, although $D \perp I$, we have $(D \not\perp I \mid G)$. This can be done by showing that $\mathbb{P}(D = d, I = i|G = g) \neq \mathbb{P}(D = d|G = g)\mathbb{P}(I = i|G = g)$ for all values of d, i, and g. For that, calculate the joint conditional distribution of (D, I) given G and compare with the product of the conditional marginal distributions of $(D|G)$ and $(I|G)$.

6. More subtle effect of v-structure: L also affects (D, I). Show numerically that $(D \not\perp I \mid L)$.

7. Remember the rule we established about d-separation and conditional independence. Based on that rule, answer V ou F to the following stochastic independence statements based on the Student BN:

$$D \perp L, \quad I \perp L, \quad S \perp L, \quad I \perp L, \quad S \perp G, \quad D \perp S$$

   the following conditional independence statements:

$$(D \perp L|I), \quad (I \perp L|S), \quad (I \perp L|G), \quad (S \perp L|I), \quad (I \perp D|G), \quad (I \perp D|L), \quad (S \perp G|L), \quad (D \perp S|L)$$

   and these additional conditional independence statements

$$(D \perp L|I, S), \quad (I \perp L|G, S), \quad (S \perp L|I, G), \quad (I \perp D|L, S), \quad (S \perp G|L, D), \quad (D \perp S|L, I)$$

8. Consider the values $\mathbb{P}(G = g^0 \mid I)$. Is this a probability distribution? If so, over what possible values?

9. Which is valid: either $\mathbb{P}(D = d, L = l|G = g) = \mathbb{P}(D = d)\mathbb{P}(L = l)$ or $\mathbb{P}(D = d, L = l|G = g) = \mathbb{P}(D = d|G = g)\mathbb{P}(L = l|G = g)$.

The next questions are general, not specific on the Student BN.

6. Consider a chain BN $X \to Z \to Y$ between three random variables. Give a numerical counter-example to show both, that $X \not\perp Y$ and that $(X \not\perp Z \mid Y)$.

7. Consider a common cause BN $X \to Z \leftarrow Y$ between three random variables. Give a counter-example to show both, that $X \not\perp Y$ and that $(X \not\perp Z \mid Y)$.

8. Given three random variables X, Y, and Z, is it valid to state that is always true that

$$\mathbb{P}(X = x, Y = y|Z = z) = \mathbb{P}(X = x|Z = z)\mathbb{P}(Y = y|Z = z)$$

9. Using your best knowledge of the problem, specify a BN for each one of the following situations. You need to identify the variables, specify their possible values (you can assume they are discrete random variables) and select a set of edges for the DAG. You do NOT need to specify the CPTs.

   - The traffic in the Antônio Carlos Avenue is classified as intense or normal. Rain or the occurrence of soccer games affect substantially the traffic.
   - To establish the premium paid by car insurance clients, the insurance companies classify the risk associated with each potential insured. Using historical databases, they have the number of claims in the last year each client was involved with. They also have the age, sex, marital status of the clients, as well as her educational level, the car model, the presence of anti-theft equipment, and the presence of airbag.
   - The Solomon-Wynne experiment on dogs were run to understand how they learn to execute a simple task: avoid an electric shock. A dog is put in a compartment, the lights are turned out and a barrier is raised. Ten seconds later an electric shock is applied. Each dog was subjected to 25 such trials. They always start receiving a shock. All dogs tested eventually learned before reaching the last trial. For each dog, the results at the i-trial is recorded as success $(Y_i = 1)$ if the dog jumps the barrier before the shock occurs, or failure $(Y_i = 0)$ otherwise.

Tabela 1: Seizure counts

Id	Y_1	Y_2	Y_3	Y_4	Trt	Base	Age
1	5	3	3	3	0	11	31
2	3	5	3	3	0	11	30
3	2	4	0	5	0	6	25
4	4	4	1	4	0	8	36

The interest was to verify how this learning occurred: by sudden insight? by cumulative learning from successes and failures? More than 10 models were considered. One of them is as follows. Assume a dog learns from previous trials, with the probability of success π_i at the i-th trial depending on the number of previous shocks and the number of previous avoidances. In particular, it has been suggested that $\pi_1 = 1$ and, for $i \geq 2$, we have $\pi_i = \alpha^{s_i} \beta^{i-s_i}$ where $s_i = \sum_{k=1}^{i} Y_i$, the total number of successes at the i-th trial. The parameters α and β are real numbers in the interval $(0, 1)$.

Write π_{i+1} as a function of π_i and Y_{i+1} and propose a BN.

- A dataset was obtained from 59 patients participating in a randomised trial of anti-convulsant therapy for epilepsy. The data was organized in a table with 59 rows, one for each patient, and 8 columns. The first column has the patient id. The next four columns record the number of seizures in four consecutive periods of two weeks each. The sixth column is a binary indicator showing either the patient received the anti-convulsant therapy or not. The seventh column shows a baseline seizure counts in the 8 weeks previous to the start of the study. The eight column is the patient age, in years. The first four rows are shown in Table 1.

  Establish a BN to describe the relationship of the variables for one single patient.

- Binary dose-response data were collected in which the numbers of beetles killed after 5 hour exposure to carbon disulphide at 8 different concentrations are recorded. Let n_i be the number of beetles exposed at the i-th concentration level and Y_i be the random variable counting the number of beetles killed.