Manuel Haußmann and Ullrich Köthe
ullrich.koethe@iwr.uni-heidelberg.de

Exercise 6

Deadline: 4.07.2018

Regulations

Please create a PDF file `probabilities.pdf` for your written answers and a Jupyter notebook `boys.ipynb` for your code (export the notebook into `boys.html` as well). Zip all files into a single archive with naming convention (sorted alphabetically by last names)

`lastname1-firstname1_lastname2-firstname2_exercise05.zip`

or (if you work in a team of three)

`lastname1-firstname1_lastname2-firstname2_lastname3-firstname3_exercise05.zip`

and upload it to Moodle before the given deadline. We will give zero points if your zip-file does not conform to the naming convention.

1 Conditional Independence (8 Points)

1. Let A, B, C three binary random variables. Prove by counterexample that

$$A \perp B \,|\, C$$

does not imply

$$A \perp B.$$

Hint: To construct the counterexample, create two tables for the conditional probabilities (one for *C=true* and one for *C=false*) that conform to the conditional independence assumptions. Then marginalize the tables over C and show that unconditional independence does not hold.

2 Boy Problem

2.1 Theoretical Calculation of the Probabilities (8 Points)

You know that Alice has two children which are not twins. What is the probability that both are boys? Compute this probability when

1. you have no additional information.
2. you meet Alice with one of her children who is a boy. Interpret this event as "At least one of Alice's children is a boy."
3. you meet Alice with one of her children who is a boy, and Alice says: "This is my first-born."
4. you meet Alice with one of her children who is a boy, and Alice says: "He was born on a Sunday."
5. you meet Alice with one of her children who is a boy, and Alice says: "Today is his birthday."

We derived the probabilities p1, p2, p3 in the lecture (repeat the derivation if you haven't been there). Prove that the probability in cases (4) and (5) is given by

$$p = \frac{2C-1}{4C-1}$$

where C is the number of states of the additional feature (that is, C=7 for the day-of-week in (4) and C=365 for the birthday in (5))

Note: Assume that having a boy or a girl is equally likely. Also all birthday dates are equally likely.

2.2 Numerical Evaluation (6 Points)

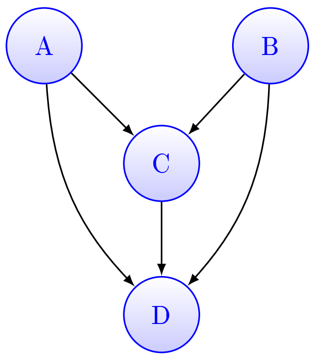
Write a program that uniformly draws N = 10000000 samples of

- sex
- day of birth
- date of birth

for each child. For every information that Alice gave you, create a subset of samples that match the condition (rejection sampling). Count the rate of (boy,boy) events in each subset and compare it with your results in section 2.1.

2.3 Belief Propagation with the sum-product algorithm (8 Points)

Here, we interpret the Boy Problem as a factor graph. We introduce the random variables A and B which indicate the sex of each of Alice's children with $P(A = 1 = \text{"}boy\text{"}) = P(B = 1 = \text{"}boy\text{"}) = 0.5$ and $C = A \vee B$ and $D = A \wedge B$. The following figure expresses this problem as a belief network:



Now consider that you meet Alice and she tells you that one of her children is a boy ($C = 1$).

Tasks:

- Draw the corresponding factor graph of the belief network shown.
- Draw the modified factor graph after you learn that Alice has at least one boy, i.e. $C = 1$.
- Which graph property has changed such that we can apply belief propagation?
- In the new graph calculate $P(A|C = 1)$ and $P(D|C = 1)$:
  - For each factor in the factor graph: Create a table that displays the factor value for all state combinations of A, B, C, D (**Note**: You can drop the states from the tables that contradict the evidence).
  - Apply the sum-product algorithm on paper (4 rounds should be sufficient).

Hint: You do not need to compute the messages to node B since they are identical to those to node A due to symmetry.

3 Weather forecast with a Markov Chain (10 Points)

We want to do a weather forecast using a simple model that classifies each day t_n as $t_n \in \{\text{Cloudy, Rainy, Sunny}\}$. Assume that each forecast only depends on the weather of the day before. The transition probabilities from day $n-1$ to day n are given as:

$$
\begin{aligned}
P(t_n = \text{Rain} \mid t_{n-1} = \text{Rain}) &= \frac{1}{5} \\
P(t_n = \text{Sunny} \mid t_{n-1} = \text{Rain}) &= \frac{1}{10} \\
P(t_n = \text{Rain} \mid t_{n-1} = \text{Sunny}) &= \frac{2}{5} \\
P(t_n = \text{Cloudy} \mid t_{n-1} = \text{Sunny}) &= \frac{3}{10} \\
P(t_n = \text{Cloudy} \mid t_{n-1} = \text{Cloudy}) &= \frac{3}{10} \\
P(t_n = \text{Rain} \mid t_{n-1} = \text{Cloudy}) &= \frac{1}{5}
\end{aligned}
$$

Tasks:

a) Derive the missing conditional probabilities.

b) Draw a directed graph with 3 nodes (Rainy,Sunny,Cloudy) and connect them with the corresponding transition probability.

c) For a given initial probability vector $\vec{p}(t_0) = \begin{pmatrix} p(t_0 = \text{Rainy}) \\ p(t_0 = \text{Cloudy}) \\ p(t_0 = \text{Sunny}) \end{pmatrix}$ derive a formula for the probabilities after one day $\vec{p}(t_1)$.

d) Cast this formula into a matrix multiplication $\vec{p}(t_1) = M\vec{p}(t_0)$ (**Hint:** don't confuse M with its transpose) and derive the numerical value for $\vec{p}(t_1)$ from the given transition probabilities M and

$$\vec{p}(t_0) = \begin{pmatrix} 0.5 \\ 0.25 \\ 0.25 \end{pmatrix}.$$

e) Generalize the formula for a forecast after n days and compute $\vec{p}(t_{100})$ numerically.

f) Find the probabilities for Rain, Sun and Clouds after infinitely many days ($\lim n \to \infty$). This is called the steady state. **Hint:** Solve the linear equations $\vec{p}(t_n) = \vec{p}(t_{n+1})$ under the constraint $|\,\vec{p}(t_n)\,|_1 = 1$