Exercises for Probabilistic Graphical Models
Sheet No.1

Bernt Schiele, Bjoern Andres, Eldar Insafutdinov, Evgeny Levinkov

Due Date: 17th November

Hand in: during the lecture or in office 629 (Eldar) or 621 (Evgeny) or by email to (eldar at mpi-inf mpg de)

You should specify your first and last name as well matriculation number on submission.

1 Probabilities

Points: 8

In this excercise, you will prove some basic, but very important rules in probability theory.

1. For any two events E_1 and E_2, prove

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2) \tag{1}$$

   what if E_1 and E_2 are two disjoint events?

2. (Bayes' law) Let $E_1, \ldots, E_n$ be mutually disjoint events in the probability space Ω such that $\Omega = \bigcup_{i=1}^{n} E_i$. Then for any event B in the same space Ω show that

$$p(E_j \mid B) = \frac{p(E_j \cap B)}{p(B)} = \frac{p(B \mid E_j)p(E_j)}{\sum_{i=1}^{n} p(B \mid E_i)p(E_i)} \tag{2}$$

   Hint: you may first want to prove the "Law of total probability" for the same events, i.e.

$$p(B) = \sum_{i=1}^{n} p(B \cap E_i) = \sum_{i=1}^{n} p(B \mid E_i)p(E_i) \tag{3}$$

3. (Linearity of expectation) For any finite collection of discrete random variables $X_1, \ldots, X_n$ with finite expectations, show that

$$\mathbb{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i] \tag{4}$$

4. Let X, Y, Z be three disjoint subsets of random variables. We say X and Y are conditionally independent given Z if and only if

$$p_{X,Y|Z}(x, y \mid z) = p_{X|Z}(x \mid z) p_{Y|Z}(y \mid z) \tag{5}$$

Show that X and Y are conditionally independent given Z if and only if the joint distribution for the three subsets of random variables factors in the following form:

$$p_{X,Y,Z}(x, y, z) = h(x, z) g(y, z) \tag{6}$$

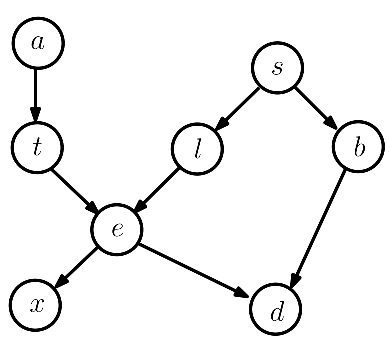
(Be careful to prove both directions!)

2 Complexity analysis

Points: 6

1. Consider the three random variables X, Y, Z all of which are binary.
   - How many parameters do you need in general to fully specify the joint distribution $p(x, y, z)$?
   - How many parameters are needed if the distribution does factorize in $p(x, y, z) = p(x \mid y)p(y \mid z)p(z)$?
   - How many parameters do you need, if the variables are not binary but can take values in $\{1, 2, \ldots, N\}$; consider both previous cases.
   - How many parameters do you need to specify a distribution over all 8bit gray-scale images of size 1000×1000 pixels? There are random variables $x_1, x_2, \ldots, x_{1M}$ with $x_i \in \{0, \ldots, 255\}$ for $i = 1, \ldots M$.
   - Do you have an idea of how to represent the distribution more compactly? Provide number of parameters needed by your method.

3 Chest Clinic Network



The chest clinic network above concerns the diagnosis of lung disease (tuberculosis,lung cancer, or both, or neiter). In this model a visit to asia is assumed to increase the probability of lung cancer. We have the following binary variables.

x	positive X-ray
d	Dyspnea (shortness of breath)
e	Either Tuberculosis or Lung Cancer
t	Tuberculosis
l	Lung cancer
b	Bronchitis
a	Visited Asia
s	Smoker

1. (**Points: 1**) Write down the factorization of the distribution implied by the graph.
2. (**Points: 4**) Are the following independence statements implied by the graph? (And how do you conclude this?)
   (a) tuberculosis$\perp\!\!\!\perp$smoking|shortness of breath
   (b) lung cancer$\perp\!\!\!\perp$bronchitis|smoking
   (c) visit to Asia$\perp\!\!\!\perp$smoking|lung cancer
   (d) visit to Asia$\perp\!\!\!\perp$smoking|lung cancer,shortness of breath
3. (**Bonus Points: 3**) Calculate by hand the values for $p(d)$. The Conditional Probability Table (CPT) is:

$$
\begin{array}{llllll}
p(a=1) & = & 0.01, & p(s=1) & = & 0.5 \\
p(t=1 \mid a=1) & = & 0.05, & p(t=1 \mid a=0) & = & 0.01 \\
p(l=1 \mid s=1) & = & 0.1, & p(l=1 \mid s=0) & = & 0.01 \\
p(b=1 \mid s=1) & = & 0.6, & p(b=1 \mid s=0) & = & 0.3 \\
p(x=1 \mid e=1) & = & 0.98, & p(x=1 \mid e=0) & = & 0.05 \\
p(d=1 \mid e=1, b=1) & = & 0.9, & p(d=1 \mid e=1, b=0) & = & 0.7 \\
p(d=1 \mid e=0, b=1) & = & 0.8, & p(d=1 \mid e=0, b=0) & = & 0.1
\end{array}
$$

and

$$
p(e=1 \mid t, l) = \begin{cases} 0 & t=0 \wedge l=0, \\ 1 & \text{otherwise.} \end{cases}
$$