CS 76140 - Spring 2017 Olga Vitek

Homework 1

Each part of the problems 5 points
Due on Blackboard before 10am on Tuesday January 24, 2016.

Note: Use any tool of your choice (including Word, latex, Markdown, or pencil+paper) to prepare the solutions. The answers should be easy to find and grade. Unreadable hand-written solutions will be given 0 points, at the grader's discretion, regardless of the correctness of the answer. For each problem, use the appropriate notation for random variables, probabilities etc. State the full formula in addition to the numerical conclusions.

1. Childhood lead poisoning is a public health concern in the United States. In a certain population, one child in 8 has a high blood lead level. In a randomly selected group of 16 children from the population, provide the probability for the following statements. Show your calculations.
   (a) None has high blood lead level?
   (b) One child has high blood lead level?
   (c) Two children have high blood lead level?
   (d) Three or more have high blood lead level?
2. The hight of corn plants follows a Normal distribution with mean 145cm and standard deviation 22cm.
   (a) What percentage of plants are between 135cm and 155cm tall?
   (b) If we choose a random sample of 16 plants, what is the probability that the mean hight would be between 135cm and 155cm tall?
   (c) If we choose a random sample of 32 plants, what is the probability that the mean hight would be between 135cm and 155cm tall?
3. In a study of behavioral asymmetries, 2,391 individuals were asked which hand they preferred to use to write, and which foot they prefer to use to kick a ball. The results are as follows:

Right hand		Left hand		Total
Right foot	Left foot	Right foot	Left foot	
2,012	142	121	116	2,391

   Are hand and foot preferences independent?

Exercises from K. Murphy:

Exercise 2.1 Probabilities are sensitive to the form of the question that was used to generate the answer

(Source: Minka.) My neighbor has two children. Assuming that the gender of a child is like a coin flip, it is most likely, a priori, that my neighbor has one boy and one girl, with probability 1/2. The other possibilities—two boys or two girls—have probabilities 1/4 and 1/4.

a. Suppose I ask him whether he has any boys, and he says yes. What is the probability that one child is a girl?
b. Suppose instead that I happen to see one of his children run by, and it is a boy. What is the probability that the other child is a girl?

Exercise 2.2 Legal reasoning

(Source: Peter Lee.) Suppose a crime has been committed. Blood is found at the scene for which there is

a. The prosecutor claims: "There is a 1% chance that the defendant would have the crime blood type if he were innocent. Thus there is a 99% chance that he is guilty". This is known as the **prosecutor's fallacy**. What is wrong with this argument?
b. The defender claims: "The crime occurred in a city of 800,000 people. The blood type would be found in approximately 8000 people. The evidence has provided a probability of just 1 in 8000 that the defendant is guilty, and thus has no relevance." This is known as the **defender's fallacy**. What is wrong with this argument?

Exercise 2.4 Bayes rule for medical diagnosis

(Source: Koller.) After your yearly checkup, the doctor has bad news and good news. The bad news is that you tested positive for a serious disease, and that the test is 99% accurate (i.e., the probability of testing positive given that you have the disease is 0.99, as is the probability of testing negative given that you don't have the disease). The good news is that this is a rare disease, striking only one in 10,000 people. What are the chances that you actually have the disease? (Show your calculations as well as giving the final result.)

Exercise 2.5 The Monty Hall problem

(Source: Mackay.) On a game show, a contestant is told the rules as follows:

> There are three doors, labelled 1, 2, 3. A single prize has been hidden behind one of them. You get to select one door. Initially your chosen door will *not* be opened. Instead, the gameshow host will open one of the other two doors, and *he will do so in such a way as not to reveal the prize.* For example, if you first choose door 1, he will then open one of doors 2 and 3, and it is guaranteed that he will choose which one to open so that the prize will not be revealed.
>
> At this point, you will be given a fresh choice of door: you can either stick with your first choice, or you can switch to the other closed door. All the doors will then be opened and you will receive whatever is behind your final choice of door.

Imagine that the contestant chooses door 1 first; then the gameshow host opens door 3, revealing nothing behind the door, as promised. Should the contestant (a) stick with door 1, or (b) switch to door 2, or (c) does it make no difference? You may assume that initially, the prize is equally likely to be behind any of the 3 doors. Hint: use Bayes rule.

Exercise 2.12 Expressing mutual information in terms of entropies

Show that

$$I(X,Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \tag{2.136}$$