Probabilistic Models and Inference
Autumn 2024

Tutorial: Week 1

Exercises

(Original from David Sontag's lecture)

1. A fair coin is tossed 4 times. Define X to be the number of heads in the first 2 tosses, and Y to be the number of heads in all 4 tosses.
   - Calculate the table of the joint probability $p(X, Y)$.
   - Calculate the tables of marginal probabilities $p(X)$ and $p(Y)$.
   - Calculate the tables of conditional probabilities $p(X|Y)$ and $p(Y|X)$.
   - What is the distribution of $Z = Y - X$?

2. You go for your yearly checkup and have several lab tests performed. A week later your doctor calls you and says she has good and bad news. The bad news is that you tested positive for a marker of 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 do not have the disease). The good news is that this is a rare disease, striking only 1 in 20,000 people. Why is it good news that the disease is rare? What are the chances that you actually have the disease?

3. Show that the statement

   $$p(A, B|C) = p(A|C)p(B|C)$$

   is equivalent to the statement

   $$p(A|B, C) = p(A|C)$$

   and also to

   $$p(B|A, C) = p(B|C)$$

   (you need to show both directions, i.e., that each statement implies the other).

4. This problem investigates the way in which conditional independence relationships affect the amount of information needed for probabilistic calculations. H, E_1, and E_2 are random variables and the notation $p(H)$ refers to the probability distribution for H, i.e., one number for every $h \in \mathrm{Val}(H)$.

   a) Suppose we wish to calculate $p(H|E_1, E_2)$, and we have no conditional independence information. Which of the following sets of numbers are sufficient for the calculation?

      i. $p(E_1, E_2), p(H), p(E_1|H), p(E_2|H)$.

      ii. $p(E_1, E_2), p(H), p(E_1, E_2|H)$.

      iii. $p(E_1|H), p(E_2|H), p(H)$.

      Provide justification for your answer.

   b) Suppose we know that E_1 and E_2 are conditionally independent given H. Now which of the above three sets are sufficient? Explain why.