Theory Assignment 4 – Basic Probability, Computing and Statistics

Fall 2015, Master of Logic, University of Amsterdam

Submission deadline: Monday, September 28th, 2015, 9 a.m.

Cooperation Cooperation among students for both theory and programming exercises is strongly encouraged. However, after this discussion phase, every student writes down and submits his/her own individual solution.

Guidelines You may pick and choose **4 exercises from exercise type I**, as well as **1 from exercise type II** for submission, i.e. you need to submit **a total of 5 exercises** to be able to get all points. Numbered exercises with an exclamation mark are supposed to be a bit harder and you may challenge yourself by trying to solve them.

In the directory of your private url there is folder called 'theory_submissions'. Please upload your submission there. Your submission should be a PDF-document (use a scanner for handwritten documents!) entitled *AssignmentX_yourStudentNumber.pdf*, where X is the number of the assignment and *yourStudentNumber* is your student number. If your submission does not comply with this format, we will deduct 1 point. For each day that your submission is late, we deduct 2 points. N.B.: If multiple files are submitted for a single assignment before the deadline, the latest version will be graded.

If you have any question about the homework or if you need help, do not hesitate to contact Thomas.

Exercises

Type I [4 exercises: 2 points per exercise]

1. An insurance company divides clients into two groups: cautious and accident prone drivers. Based on their statistics, the probability that an accident prone client will actually be involved in an accident within a one year period is .4, whereas the probability for cautious drivers is .2. Under the assumption that 40% of the population is accident prone, calculate the probability that a new client will have an accident within a year of purchasing an insurance.

2. Many multiple-choice tests do not deduct points for wrong answers. Therefore, should one not know the answer, guessing is a recommendable strategy. Let k be the probability that a student knows the answer to a multiple-choice question, and $1 - k$ that he guesses instead. Further, let $\frac{1}{c}$ be the probability of guessing correctly, where c is the number of alternatives in the multiple-choice question. Give and interpret the probability that a student knew the answer to a question given that it was answered correctly as a function of k, $k \in [0, 1]$, and $c = 5$ and $c = 100$.

3. A friend has 3 cards, identical in form but not in color: the first is green from both sides, the second is yellow from both sides, and the third has one yellow and one green side. She mixes the three cards, selects one at random, and puts it down on the table. The card shows a green side. Would you rather bet money on the card's other side being yellow, green, or either? Indicate the relevant computations and reasoning used to arrive at this conclusion.

4. A factory has 11 automatic machines that drop their produced goods into 11 unlabeled boxes. The production of the goods is highly complex and prone to error. As a consequence, ten of the machines produce one in six defective goods. Matters are worse for the eleventh machine as it produces one in three defective goods. Each of the machines' boxes is shipped to quality control with 12 goods inside. Your job at quality control is to decide whether the goods of a box were produced by the bad eleventh machine or not. To save costs only some of the box's contents are inspected. Let G denote a 'good' product and B a 'bad' one. Compute the probability of having a box from the eleventh machine when, after subsequent inspection, witnessing

   (i) B (the first product was bad);

   (ii) BBG (the first two products were bad, the third good);

   (iii) $BBGBBB$ (the first two were bad, the third good, the remainder bad).

5. Consider the premise 'An observation in line with a hypothesis supports the hypothesis', and the claim 'It follows from the hypothesis that "all swans are white" that "all non-white things are non-swans", and this is supported by the observation of a black tie.'. Qualify this statement using a probabilistic argument [Hint: examples with concrete numeric values may help].

Type II [1 exercise: 2 points per exercise]

1. Let E_1, E_2 and H be events. Consider the following three sets of values

   (a) $P(E_1 \cap E_2), P(H), P(E_1|H)$, and $P(E_2|H)$;

   (b) $P(E_1 \cap E_2), P(H)$, and $P(E_1 \cap E_2|H)$;

   (c) $P(E_1|H), P(E_2|H)$, and $P(H)$.

   Answer:

   (i) Which of these sets of values are sufficient to compute $P(H|E_1 \cap E_2)$? Justify your answer by showing for each set either how to calcuate the desired outcome, or explain why this is not possible.

   (ii) Let $P(E_1 \cap E_2|H) = P(E_1|H)P(E_2|H)$ and answer the question in (i) with this new information. Justify your answers as above.

2. (i) Show that

$$\frac{P(H|E)}{P(I|E)} = \frac{P(E|H)P(H)}{P(E|I)P(I)}$$

   (ii) Let H be four times more likely than I prior to new evidence and the new evidence be twice as likely under I than H. Which hypothesis is more likely after observing the evidence?