Assignment 1

Machine Learning Applications, Fall 2019

- These are traditional textbook style exercises. In all the questions that involve calculations, you are required to show all your work. That is, you need to write the steps that you made in order to get to the solution (see latex template and relative pdf in Sulis for an example).

- Solutions should be uploaded to Sulis before the due date; they should be each a single PDF document, and additional files will not be considered.

- Students may (and should) collaboratively discuss the assignments; however, I expect each student to write and upload their own solutions. Please write your full name and the names of the students (if any) you discussed the assignment with at the top of your solutions.

Exercise 1 (points 0.5)

We toss a fair coin 3 times.

a Write down the possibility space, if we record the exact sequences of Heads($=1$) and Tails ($=0$)

b Write down the possibility space, if we record only the total number of Heads.

Exercise 2 (points 0.5)

Assign a probability to all elements (elementary events) of the possibility spaces of Exercise 1.(a) and 1.(b).

Exercise 3 (points 0.5)

We randomly extract 3 balls from an urn with 365 balls, numbered $1, \ldots, 365$.

a If we put each ball back into the urn before we draw the next, how many possible outcomes of the experiment are there?

b Answer the same question as above, but now if we do not put the balls back.

c Calculate the probability that we draw 3 times the same ball in case (a).

Exercise 4 (points 0.5)

Consider the experiment of throwing 2 fair dice.

a Find the probability that both dice show the same face.

b Find the same probability, given you know that the sum of the dice is not greater than 4.

Exercise 5 (points 1)

Consider the experiment of throwing 2 dice. Denote with x the outcome for the first dice and y the ourcome for the second dice. Assume that the joint probability mass function of x, y is:

$$p(x = i, y = i) = 0 \text{ and } p(x = i, y = j) = \frac{1}{30}$$

for $i, j = 1, 2, 3, 4, 5, 6$ and $i \neq j$.

a Compute the marginal probability mass functions $p(x)$ and $p(y)$.

b Are the two variables x and y independent?

c By applying Bayes' rule, compute the probability that $x + y \geq 11$ given that you know that x is even.

Exercise 6 (points 1)

In a binary transmission channel, the bit 1 is transmitted with probability 2/3 and the bit 0 with probability 1/3. The channel is noisy so the conditional probability of receiving a 1 when a 1 was sent is 0.95, the conditional probability of receiving a 0 when a 0 was sent is 0.90. Given that a 1 is received, what is the probability that a 1 was transmitted?

Exercise 7 (points 1)

Let x be a discrete variable that can assume values in the possibility space $\Omega = \{x_1, \ldots, x_n\}$. Using the definition of expectations of a function $f(x)$ of x:

$$E[f(x)] = \sum_{i=1}^{n} f(x_i) p(x_i)$$

where $p(x_i)$ is the probability mass function of x.

Prove that

a $E[c\,x] = c\,E[x]$ where c is a constant.

b $E[(x - E[x])^2] = E[x^2] - E[x]^2$.

Exercise 8 (points 1)

Suppose the probability that we are flu is $p(\text{Flu}) = 0.05$, the probability that we have High-temperature when we are flu is $p(\text{High-temperature}|\text{Flu}) = 0.9$

and the probability that we have a high temperature when we are not flu (false alarm) is $p(\text{High-temperature}|\neg\text{Flu}) = 0.2$, where $\neg$ denotes not.

Assume we have evidence that a person has a high-temperature, what is the probability that this person is Flu?

Exercise 9 (points 2)

Consider the last exercise. We have now a thermometer whose rate of false negative reading is 5% and false positive reading is 14%, that is,

$$p(\text{HighTherm} = True|\text{HighTemp} = True) = 0.95$$
$$p(\text{HighTherm} = True|\text{HighTemp} = False) = 0.15$$

Assume that Flu and HighTherm are independent given that the person has HighTemp.

a The person is Flu and Thermometer suggests HighTemp, what is the probability that the person has a High temperature?

b Thermometer suggests HighTemp, what is the probability of the person being Flu?

Exercise 10 (points 2)

Two men have left traces of their own DNA at the scene of a crime. A suspect, Jack, is tested and found to have a gene "XXZ'. The DNA of the two traces are found to be of type "XXZ" (a common type in the local population, having frequency 60%) and type "XXW" (a rare type, with frequency 1%). Does this data (type "XXZ" and "XXW" that were found at the scene) give evidence in favor of the hypothesis that Jack was one of the two people present at the scene of the crime?