UNIVERSITÉ CÔTE D'AZUR

Cédric Richard

Quantitative Measure of Information
Part II

Exercise

Let X be a discrete random variable that can take n possible values, and Y a discrete random variable uniformly distributed that can take n possible values. Throughout the exercise, no assumption will be made on the distribution of X.

1. Calcule the maximum entropy that can be reached by X, and specify the case in which this value would be obtained.

2. Calcule the entropy of Y.

3. In the case of $n = m$, rank in ascending order the following quantities: $H(Y)$, 0, $H(X)$, $H(X;Y)$, $H(X) + H(Y)$, $H(X|Y)$. Justify each step of your ranking.

4. Explain cases of equality in the ranking proposed in 3., i.e., give for each inequality the cases in which it becomes an equality.

5. We now consider the case where the joint distribution $P(X = x_i; Y = y_j)$ is given by the table below:

$P(X;Y)$	y_1	y_2	y_3	y_4
x_1	1/24	1/12	1/6	1/24
x_2	1/6	1/8	1/24	1/6
x_3	1/24	1/24	1/24	1/24

Check that Y is uniformaly distributed. Calculate $I(X;Y)$.

Problem

For a given region, the forecasts of a meteorologist are divided according to their relative frequencies given by the table below. The columns correspond to the actual weather, which is represented by the random variable T, which takes values 0 or 1 depending on whether the weather is rainy or sunny, respectively. The rows correspond to the meteorologist's forecast, identified by the random variable M, also with values in $\{0.1\}$ depending on whether he had planned a rainy weather (0) or a sunny weather (1).

$P(M = i, T = j)$	sunny weather ($T = 1$)	rainy weather ($T = 0$)
sunny weather ($M = 1$)	5/8	1/16
rainy weather ($M = 0$)	3/16	1/8

1. Calculate the probabilities $P(M = i)$ and $P(T = j)$, with $i, j \in \{0, 1\}$.

2. Show that the meteorologist is wrong once in 4 times.

3. One student says that by always forecasting sunny weather, he makes fewer mistakes than the meteorologist does. Check this assertion.

4. Let E be the random variable representing the student's prediction. As for T and M, random variable E takes values in $\{0.1\}$. Calculate $I(E;T)$.

5. Calculate $I(M;T)$.

6. Comparing $I(M;T)$ to $I(E;T)$, what Information Theory shows on the meteorologist's forecast and that of the student?

7. The student claims to have found a revolutionary method of predicting the weather. Its revised performance are provided in the table above. As before, the rows correspond to the forecast, and the columns to the actual weather.

$P(E = i, T = j)$	sunny weather $(T = 1)$	rainy weather $(T = 0)$
sunny weather $(E = 1)$	403/512	93/512
rainy weather $(E = 0)$	13/512	3/512

Calculate the probabilities $P(E = 0)$ and $P(E = 1)$.

8. Compare $P(E = i, T = j)$ and $P(E = i)\, P(T = j)$, for all i, $j \in \{0, 1\}$. Conclude.

9. We wish to store T by using a binary coding. Using Shannon's first theorem, give the minimum average memory space required to store T, in bits per realization of T.

11. Redo the previous calculation in the case of M. Calculate the minimum memory space required to store M and T separately, in bits per realization of (M, T)?

12. Calculate the minimum memory space required to store M and T jointly, in bits per realization of (M, T)?

13. Interpret the difference between results of the 2 previous questions.

14. Propose Huffman coding to jointly encode M and T.

15. Calculate the average length of words $\overline{n}$ of the binary code found in the previous question. What double inequality is satisfied by $\overline{n}$?