

Exercise of Supervised Learning: Information Theory Part 2

Yawei Li

yawei.li@stat.uni-muenchen.de

February 2, 2024

Exercise 1: Entropy

A fair die is rolled at the same time as a fair coin is tossed. Let A be the number on the upper surface of the dice and let B describe the outcome of the coin toss, where

$$B = \begin{cases} 1, & \text{head,} \\ 0 & \text{tail.} \end{cases}$$

Two random variables X and Y are given by $X = A + B$ and $Y = A - B$, respectively.

(a) Calculate the entropies $H(X)$ and $H(Y)$, the conditional entropies $H(Y|X)$ and $H(X|Y)$, the joint entropy $H(X, Y)$ and the mutual information $I(X; Y)$.

Exercise 1 (a)

(a) Calculate the entropies $H(X)$ and $H(Y)$, the conditional entropies $H(Y|X)$ and $H(X|Y)$, the joint entropy $H(X, Y)$ and the mutual information $I(X; Y)$.

Solution: Let a, b, x , and y denote the realizations of A, B, X and Y , respectively.

Note that each event (a, b) is associated with **exactly one** event (x, y) . Why?

Because given a pair (x, y) , $a = 0.5(x + y)$ and $b = 0.5(x - y)$ are unique. So the joint probability

$$p_{AB}(a, b) = p_{XY}(x, y) = \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12}$$

So, the joint entropy:

$$\begin{aligned} H(X, Y) &= - \sum_{x,y} p_{X,Y}(x, y) \log_2 p_{X,Y}(x, y) \\ &= -12 \cdot \frac{1}{12} \log_2 \frac{1}{12} \\ &= 2 + \log_2 3 \end{aligned}$$

Exercise 1 (a): Continued

x	Events (a, b)	$p_X(x)$
1	(1, 0)	1/12
2	(2, 0), (1, 1)	1/6
3	(3, 0), (2, 1)	1/6
4	(4, 0), (3, 1)	1/6
5	(5, 0), (4, 1)	1/6
6	(6, 0), (5, 1)	1/6
7	(6, 1)	1/12

y	Events (a, b)	$p_Y(y)$
0	(1, 1)	1/12
1	(1, 0), (2, 1)	1/6
2	(2, 0), (3, 1)	1/6
3	(3, 0), (4, 1)	1/6
4	(4, 0), (5, 1)	1/6
5	(4, 0), (6, 1)	1/6
6	(6, 0)	1/12

$$H(X) = \sum_x p_X(x) \log_2 p_X(x) = \frac{7}{6} + \log_2 3.$$

$$H(Y) = \sum_y p_Y(y) \log_2 p_Y(y) = \frac{7}{6} + \log_2 3.$$

Exercise 1 (a): Continued

Disclaimer: the basic formulas like the following ones need to be either remembered or written on your cheatsheet. They will not be given on the exam sheets.

The conditional entropies

$$H(X|Y) = H(X, Y) - H(Y) = \frac{5}{6}$$

$$H(Y|X) = H(X, Y) - H(X) = \frac{5}{6}$$

The mutual information

$$I(X; Y) = H(X) - H(X|Y) = \frac{1}{3} + \log_2 3.$$

Exercise 1 (b)

(b) Show that, for independent discrete random variables X and Y ,

$$I(X; X + Y) - I(Y; X + Y) = H(X) - H(Y).$$

Solution:

$$\begin{aligned} I(X; X + Y) - I(Y; X + Y) &= H(X + Y) - H(X + Y|X) - H(X + Y) + H(X + Y|Y) \\ &= H(X + Y|Y) - H(X + Y|X) \end{aligned}$$

Note that $p(x + y|x) = p(y|x)$, because x is given, so if we observe $x + y$, we can immediately infer the value of y . In this case where X is observed, there is an one-to-one mapping between $X + Y$ and Y .

So, $H(X + Y|X) = H(X|Y)$, and $H(X + Y|Y) = H(X|Y)$. Hence,

$$\begin{aligned} I(X; X + Y) - I(Y; X + Y) &= H(X|Y) - H(Y|X) \\ &= H(X) - H(Y) \end{aligned}$$

Exercise 2 (a)

Let X , Y and Z be three discrete random variables. The mutual information of X , Y and Z is defined as:

$$I(X; Y; Z) = \sum_x \sum_y \sum_z p(x, y, z) \log \left(\frac{p(x, y)p(x, z)p(y, z)}{p(x)p(y)p(z)p(x, y, z)} \right)$$

(a) Prove the lemma: $I(X; Y; Z) = I(X; Y) - I(X; Y|Z)$. Note that the conditional mutual information is defined as:

$$I(X; Y|Z) = \sum_x \sum_y \sum_z p(z)p(x, y|z) \log \frac{p(x, y|z)}{p(x|z)p(y|z)}$$

Hint: Starting from the right hand side of the equation in the lemma.

Exercise 2 (a): Continued

$$\begin{aligned} I(X; Y) - I(X; Y|Z) &= \sum_x \sum_y p(x, y) \log \frac{p(x, y)}{p(x)p(y)} - \sum_z \sum_x \sum_y p(z)p(x, y|z) \log \frac{p(x, y|z)}{p(x|z)p(y|z)} \\ &= \sum_x \sum_y \sum_z p(x, y, z) \log \frac{p(x, y)}{p(x)p(y)} \\ &\quad - \sum_x \sum_y \sum_z p(x, y, z) \log \frac{p(x, y|z)p(z)^2}{p(x|z)p(y|z)p(z)^2} \\ &= \sum_x \sum_y \sum_z p(x, y, z) \log \frac{p(x, y)}{p(x)p(y)} - \sum_x \sum_y \sum_z p(x, y, z) \log \frac{p(x, y, z)p(z)}{p(x, z)p(y, z)} \\ &= \sum_x \sum_y \sum_z p(x, y, z) \log \left(\frac{p(x, y)p(x, z)p(y, z)}{p(x)p(y)p(z)p(x, y, z)} \right) \\ &= I(X; Y; Z) \end{aligned}$$

Exercise 2 (b)

(b) Prove the following relation with the above lemma:

$$I(X; Y) = I(X; Y|Z) + I(Y; Z) - I(Y; Z|X).$$

Solution: Using the lemma on the conditional MI, we obtain:

$$\begin{aligned} & I(X; Y|Z) + I(Y; Z) - I(Y; Z|X) \\ &= I(X; Y) - I(X; Y; Z) + I(Y; Z) - I(Y; Z) + I(X; Y; Z) \\ &= I(X; Y). \end{aligned}$$

Exercise 3

Show the standard solution.