

Introduction to Information Theory, Fall 2019

Homework problem set #2

due September 20, 2019

Rules: Always explain your solutions carefully. You can work in groups, but must write up your solutions alone. You must submit your solutions before the Friday exercise class (either in person or by email).

1. **Entropy, essential bit content, typical set (1 point):** Let P be the probability distribution with three possible outcomes A, B, C and probabilities $P(A) = 1/2, P(B) = 1/4, P(C) = 1/4$. Let X_1, X_2 be independent and identically distributed (IID) according to P .
 - (a) Compute $H(X_1), H(X_2)$, and $H(X_1, X_2)$.
 - (b) Make a table that lists the joint probability $P(x_1, x_2)$ and the quantity $\frac{1}{2} \log \frac{1}{P(x_1, x_2)}$ for all possible outcomes x_1 and x_2 .
 - (c) Compute $H_\delta(X_1, X_2)$ for $\delta = 3/8$.
 - (d) Write down all elements of the typical set $T_{2, \varepsilon}(P)$ for $\varepsilon = 0.12345$.

2. **Subadditivity of the entropy (1 point):** The goal of this problem is to show that

$$H(X, Y) \leq H(X) + H(Y) \tag{1}$$

for any two random variables X, Y with an arbitrary joint distribution $P(x, y)$.

- (a) Can you interpret the inequality in the context of compression?
- (b) Verify the following identity, where $P(x) = \sum_y P(x, y)$ and $P(y) = \sum_x P(x, y)$ denote the marginal distributions (as always):

$$H(X) + H(Y) - H(X, Y) = \sum_{x, y} P(x, y) \log \frac{P(x, y)}{P(x)P(y)}$$

- (c) Use Jensen's inequality to prove Eq. (1).

Optional problem: Show that equality holds precisely when X and Y are independent.

3. **Lossless compression (1 point):**

In this problem, you will implement the 'universal' compression algorithm discussed in class to compress and decompress pictures. To get started, open the Python notebook at <https://colab.research.google.com/github/amsqi/iit19-homework/blob/master/02-homework.ipynb> and follow the instructions.

As last week, please submit your solution as a Python notebook or script, or as a PDF printout. You can score the maximum score if your solution produces the correct output. We will only have a closer look at your code in case of problems.