INFORMATION THEORY
Master of Logic, University of Amsterdam, 2017
TEACHER: Christian Schaffner, TA: Yfke Dulek, Alvaro Piedrafita

Homework problem set 2

Deriving the weak law of large numbers

Problem 1

(Markov's inequality) For any real non-negative random variable X and any $t > 0$, show that

$$P_X(X \geq t) \leq \frac{\mathbb{E}[X]}{t}\,.$$

Problem 2

Exhibit a random variable (which can depend on t) that achieves this inequality with equality.

Problem 3

(Chebyshev's inequality.) Let Y be a random variable with mean μ and variance σ^2. Show that for any $\varepsilon > 0$,

$$P(|Y - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}\,.$$

Hint: Define a random variable $X := (Y - \mu)^2$.

Problem 4

(The weak law of large numbers.) Let $Z_1, Z_2, ..., Z_n$ real i.i.d. random variables with mean $\mu = \mathbb{E}[Z_i]$ and variance $\sigma^2 = \mathbb{E}[(X_i - \mu)^2] < \infty$. Define the random variables $S_n = \frac{1}{n}\sum_{i=1}^n Z_i$. Show that

$$P(|S_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}\,.$$

Thus, $P(|S_n - \mu| \geq \varepsilon) \to 0$ as $n \to \infty$. This is known as the weak law of large numbers (Theorem 2.6.1 in the lecture notes).

Huffman Coding

Problem 5

Consider the binary source P_X with $P_X(0) = \frac{1}{8}$ and $P_X(1) = \frac{7}{8}$.

Design a Huffman code for P_X. Describe, in order, which source symbols you combine, and list the final codebook. What is the average codeword length?

Problem 6

Design a Huffman code for blocks of $N = 2$ bits drawn from the source P_X. Describe, in order, which source symbols you combine, and list the final codebook. What is the average codeword length?

Problem 7

Design a Huffman code for blocks of $N = 3$ bits drawn from the source P_X. Describe, in order, which source symbols you combine, and list the final codebook. What is the average codeword length?

Problem 8

For the three codes you designed ($N = 1, 2, 3$), divide the average codeword length by N, and compare these values to the optimal length, i.e., $H(X)$. What do you observe?

Problem 9

If you were asked to design a Huffman code for a block of $N = 100$ bits, what problem would you run into?

Problem 10

Consider the random variable Z with

z	1	2	3	4	5	6
$P_Z(z)$	1/10	3/10	2/10	2/10	1/10	1/10

Design a *ternary* Huffman code for Z (i.e. using an alphabet with three symbols).

Inefficiency when using the wrong code

Problem 11

Given are two distributions P_X and Q_X for the same set $\mathcal{X} = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}\}$:

x	a	b	c	d
$P_X(x)$	1/4	1/4	1/4	1/4
$Q_X(x)$	1/2	1/4	1/8	1/8

Design an optimal prefix-free code for $Q_X(x)$.

Problem 12

What is the expected codeword length of the code you just designed? (Three decimals precision)

Problem 13

What is the expected codeword length if you use this code to encode the source P_X? (Three decimals precision)

Problem 14

Show that in general, when using optimal code for any source source Q_X (which has lengths $\ell(x) = \lceil -\log Q_X(x) \rceil$ for all $x \in \mathcal{X}$) to encode another source P_X, this incurs a penalty of $D(P_X||Q_X)$ in the average description length.

More formally, prove that

$$H(P_X) + D(P_X||Q_X) \leq \mathbb{E}_{P_X}[\ell(X)] \leq H(P_X) + D(P_X||Q_X) + 1$$

for all P_X and Q_X.