Spring 2016

6.441 - Information Theory

Homework 3

Due: Thur, Feb 25, 2016 (in class)

`Prof. Y. Polyanskiy`

1 Reading (optional)

1. Read [1, Chapters 3,4]

2 Exercises

NOTE: Each exercise is 10 points. Only 3 exercises per assignment will be graded. If you submit more than 3 solved exercises please indicate which ones you want to be graded.

1 (Maximum entropy.) Prove that for any X taking values on $\mathbb{N} = \{1, 2, \ldots\}$ such that $\mathbb{E}X < \infty$,

$$H(X) \le \mathbb{E}X h\left(\frac{1}{\mathbb{E}X}\right).$$

Find the necessary and sufficient condition for equality. *Hint*: Find an appropriate Q such that RHS - LHS $= D(P_X||Q)$.

2 Show that for jointly gaussian (A, B, C)

$$I(A; C) = I(B; C) = 0 \quad \Longrightarrow \quad I(A, B; C) = 0\,. \tag{1}$$

Find a counter-example for general (A, B, C).

Prove or disprove: Implication (1) also holds for arbitrary discrete (A, B, C) under positivity condition $P_{ABC}(a, b, c) > 0\ \forall abc$.

3 (Divergence of order statistics) Given $x^n = (x_1, \ldots, x_n) \in \mathbb{R}^n$, let $x_{(1)} \le \ldots \le x_{(n)}$ denote the ordered entries. Let P, Q be distributions on $\mathbb{R}$ and $P_{X^n} = P^n, Q_{X^n} = Q^n$.

1. Prove that

$$D(P_{X_{(1)},\ldots,X_{(n)}}||Q_{X_{(1)},\ldots,X_{(n)}}) = nD(P||Q). \tag{2}$$

2. Show that

$$D(\text{Binom}(n, p)||\text{Binom}(n, q)) = nd(p||q).$$

4 *Run-length encoding* is a popular variable-length lossless compressor used in fax machines, image compression, etc. Consider compression of S^n – an i.i.d. Bern(δ) source with very small $\delta = \frac{1}{128}$ using run-length encoding: A chunk of consecutive $r \le 255$ zeros (resp. ones) is encoded into a zero (resp. one) followed by an 8-bit binary encoding of r (If there are > 255 consecutive zeros then two or more 9-bit blocks will be output). Compute the average achieved compression rate

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\ell(f(S^n)\right]$$

How does it compare with the optimal lossless compressor?

Hint: Compute the expected number of 9-bit blocks output per chunk of consecutive zeros/ones; normalize by the expected length of the chunk.

5 Recall that an entropy rate of a process $\{X_j : j = 1, \ldots\}$ is defined as follows provided the limit exists:

$$\mathcal{H} = \lim_{n\to\infty} \frac{1}{n} H(X^n).$$

Consider a 4-state Markov chain with transition probability matrix

$$\begin{bmatrix} 0.89 & 0.11 & 0 & 0 \\ 0.11 & 0.89 & 0 & 0 \\ 0 & 0 & 0.11 & 0.89 \\ 0 & 0 & 0.89 & 0.11 \end{bmatrix}$$

The distribution of the initial state is $[p, 0, 0, 1-p]$.

1. Does the entropy rate of such a Markov chain exist? If it does, find it.
2. Describe the asymptotic behavior of the optimum variable-length rate $\frac{1}{n}\ell(f^*(X_1, \ldots, X_n))$. Consider convergence in probability and in distribution.
3. Repeat with transition matrix:

$$\begin{bmatrix} 0.89 & 0.11 & 0 & 0 \\ 0.11 & 0.89 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{bmatrix}$$

6 (Elias coding) In this problem all logarithms are binary.

1. Consider the following universal compressor for natural numbers: For $x \in \mathbb{N} = \{1, 2, \ldots\}$, let $k(x)$ denote the length of its binary representation. Define its codeword $c(x)$ to be $k(x)$ zeros followed by the binary representation of x. Compute $c(10)$. Show that c is a prefix code and describe how to decode a stream of codewords.
2. Next we construct another code using the one above: Define the codeword $c'(x)$ to be $c(k(x))$ followed by the binary representation of x. Compute $c'(10)$. Show that c' is a prefix code and describe how to decode a stream of codewords.
3. Let X be a random variable on $\mathbb{N}$ whose probability mass function is decreasing. Show that $\mathbb{E}[\log(X)] \leq H(X)$.
4. Show that the average code length of c satisfies $\mathbb{E}[l(c(X))] \leq 2H(X) + 2$ `bit`.
5. Show that the average code length of c' satisfies $\mathbb{E}[c'(X)] \leq H(X) + 2\log(H(X)+1) + 3$ `bit`.

Comments: The two coding schems are known as Elias γ-codes and δ-codes.

References

[1] T. Cover and J. Thomas, *Elements of Information Theory,* Second Edition, Wiley, 2006