Information Theory SS 2025

Institut für Diskrete Mathematik (5050), TU Graz

Exercise sheet 4 (15.05.2025)

Exercise 4.1. Find an example of two different probability distributions $p \neq q$ on the set $\{0, 1\}$ that satisfy $D(p \parallel q) = D(q \parallel p)$.

Exercise 4.2. Let $X_1, \dots, X_n, Y$ be random variables. Prove the chain rule for mutual information, that is, show explicitly that

$$I((X_1, \dots, X_n); Y) = I(X_1; Y) + I(X_2; Y|X_1) + \dots + I(X_n; Y|X_1, \dots, X_{n-1})$$

Give a more general chain rule for expressions of the form $I((X_1, \dots, X_n); (Y_1, \dots, Y_m))$.

Exercise 4.3. Our goal is to identify a random object X which is distributed in $\mathcal{X}$ with some distribution p. A question Q from a set $\mathcal{Q}$ is asked at random according to distribution r. This results in a deterministic answer $A = A(X, Q)$, that is, there is a deterministic function $(x, q) \mapsto A(x, q)$ from $\mathcal{X}$ to $\mathcal{Q}$. Suppose X and Q are independent. Then $I(X; Q, A)$ is the uncertainty in X removed by the question-answer pair (Q, A).

(a) Show that $I(X; Q, A) = H(A \mid Q)$ and interpret this statement.

(b) Now suppose that two i.i.d. questions $Q_1, Q_2 \sim r$ are asked, yielding answers A_1 and A_2. Show that $I(X; Q_1, A_1, Q_2, A_2) \leq 2I(X; Q_1, A_1)$.

Exercise 4.4. The *interaction information* of three random variables X_1, X_2, X_3 is defined as

$$I(X_1\,;\,X_2\,;\,X_3) := I(X_1\,;\,X_2) - I(X_1\,;\,X_2 \mid X_3).$$

(a) Show that the interaction information is symmetric, in the sense that

$$I(X_1\,;\,X_2\,;\,X_3) = I(X_{\sigma(1)}\,;\,X_{\sigma(2)}\,;\,X_{\sigma(3)})$$

holds for any permutation σ of $\{1, 2, 3\}$.

(b) Give an example of random variables X_1, X_2, X_3 such that $I(X_1\,;\,X_2\,;\,X_3) < 0$.

Exercise 4.5. Let X, Y, Z_1, Z_2 be random variables on a common state space $\mathcal{X}$ forming the Markov triple $(X, Y, (Z_1, Z_2))$, that is,

$$p(x, y, z_1, z_2) = p(x)\,p(y \mid x)\,p(z_1, z_2 \mid y)$$

holds for all $x, y, z_1, z_2 \in \mathcal{X}$. Show that

$$I(X\,;\,Z_1) + I(X\,;\,Z_2) \leq I(X\,;\,Y) + I(Z_1\,;\,Z_2).$$

Exercise 4.6. Consider the Markov chain $(X_n)_{n\geq 0}$ on $\mathcal{X} = \{1, \dots, 7\}$ with transition matrix

$$\mathbf{P} = \begin{pmatrix} \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \end{pmatrix}.$$

(a) Draw the transition graph. Is (X_n) irreducible?

(b) Calculate $\mathbb{P}[X_n = i \mid X_0 = 6]$ for $i \in \mathcal{X}$ and $n \in \{1, 2, 3\}$.

(c) Give a stationary distribution for X_n. Is it unique?

(d) Calculate $\lim_{n\to\infty} \mathbb{P}[X_n = i \mid X_0 = 6]$ for all $i \in \mathcal{X}$.

Exercise 4.7. Give an example to show that in general, a function of a Markov chain is not necessarily again a Markov chain (i.e., find a Markov chain $(X_n)_{n\geq 0}$ with finite state space $\mathcal{X}$ and a function $f : \mathcal{X} \to \mathcal{Y}$ such that $(Y_n)_{n\geq 0}$ with $Y_n = f(X_n)$ is not a Markov chain).

Exercise 4.8. Let $(X_n)_{n\geq 0}$ be an irreducible time homogeneous Markov chain with transition matrix P and stationary initial distribution ν. Let $N \in \mathbb{N}$ and consider the stochastic process $(Y_n)_{n=0}^N$ where $Y_n = X_{N-n}$.

(a) Show that $(Y_n)_{n=0}^N$ is a Markov chain.

(b) Determine the transition matrix for the chain.

(c) Show that the reversed chain is stationary and determine its stationary distribution.