

**1. Entropy and Mutual Information (20 points)**

- (a) Let  $X$  be a discrete random variable on a space  $\mathcal{X}$  with  $|\mathcal{X}| = k < \infty$ .
- Prove that the entropy  $H(X) \geq 0$ , with equality only when  $X$  is a constant.
  - Denote by  $p$  the distribution of  $X$  and  $q$  the uniform distribution on  $\mathcal{X}$ . What is the relation between the Kullback-Leibler divergence  $D(p||q)$  and the entropy  $H(X)$  of the distribution  $p$ ?
  - Deduce a tight upper bound which depends on  $k$  for the entropy of any distribution  $p$  over  $\mathcal{X}$ .
- (b) We consider a pair of discrete random variables  $(X_1, X_2)$  defined over the finite set  $\mathcal{X}_1 \times \mathcal{X}_2$ . Let  $p_{1,2}$ ,  $p_1$  and  $p_2$  denote respectively the joint distribution, the marginal distribution of  $X_1$  and the marginal distribution of  $X_2$ . We define the mutual information between  $X_1$  and  $X_2$  as:

$$I(X_1, X_2) := \sum_{(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2} p_{1,2}(x_1, x_2) \log \frac{p_{1,2}(x_1, x_2)}{p_1(x_1)p_2(x_2)}.$$

- Manipulate the expression above to show that  $I(X_1, X_2) \geq 0$ .
- Let  $H(Z)$  be the entropy of the random variable  $Z = (X_1, X_2)$ . Show that  $I(X_1, X_2)$  can be expressed as a function of  $H(X_1)$ ,  $H(X_2)$  and  $H(Z)$ .
- What is the joint distribution  $p_{1,2}$  over  $\mathcal{X}_1 \times \mathcal{X}_2$  of maximal entropy with fixed marginals  $p_1$  and  $p_2$ ?

**2. Hidden Markov Models (80 points)**

Follow the instructions in this Colab notebook: [https://colab.research.google.com/drive/1cfWlGh9-o0L-2RThujXzW\\_VgYZn9V5M6](https://colab.research.google.com/drive/1cfWlGh9-o0L-2RThujXzW_VgYZn9V5M6)