COMP2610/COMP6261 - Information Theory

Tutorial 5: Probabilistic inequalities and Mutual Information

Young Lee and Bob Williamson
Tutors: Debashish Chakraborty and Zakaria Mhammedi

Week 6 (28th Aug- 1st Sep), Semester 2, 2017

1. Consider a discrete variable X taking on values from the set $\mathcal{X}$. Let p_i be the probability of each state, with $i = 1, \ldots, |\mathcal{X}|$. Denote the vector of probabilities by $\mathbf{p}$. We saw in lectures that the entropy of X satisfies:

$$H(X) \leq \log|\mathcal{X}|,$$

with equality if and only if $p_i = \frac{1}{|\mathcal{X}|}$ for all i, i.e. $\mathbf{p}$ is uniform. Prove the above statement using Gibbs' inequality, which says

$$\sum_{i=1}^{|\mathcal{X}|} p_i \log_2 \frac{p_i}{q_i} \geq 0$$

for any probability distributions $\mathbf{p}, \mathbf{q}$ over $|\mathcal{X}|$ outcomes, with equality if and only if $\mathbf{p} = \mathbf{q}$.

2. Let X be a discrete random variable. Show that the entropy of a function of X is less than or equal to the entropy of X by justifying the following steps:

$$\begin{aligned} H(X, g(X)) &\stackrel{(a)}{=} H(X) + H(g(X)|X) \\ &\stackrel{(b)}{=} H(X); \\ H(X, g(X)) &\stackrel{(c)}{=} H(g(X)) + H(X|g(X)) \\ &\stackrel{(d)}{\geq} H(g(X)). \end{aligned}$$

Thus $H(g(X)) \leq H(X)$.

3. Random variables X, Y, Z are said to form a Markov chain in that order (denoted by $X \to Y \to Z$) if their joint probability distribution can be written as:

$$p(X, Y, Z) = p(X) \cdot p(Y|X) \cdot p(Z|Y)$$

   (a) Suppose (X, Y, Z) forms a Markov chain. Is it possible for $I(X;Y) = I(X;Z)$? If yes, give an example of X, Y, Z where this happens. If no, explain why not.

   (b) Suppose (X, Y, Z) does *not* form a Markov chain. Is it possible for $I(X;Y) \geq I(X;Z)$? If yes, give an example of X, Y, Z where this happens. If no, explain why not.

4. If $X \to Y \to Z$, then show that

   (a) $I(X;Z) \leq I(X;Y)$

   (b) $I(X;Y|Z) \leq I(X;Y)$

5. A coin is known to land heads with probability $\frac{1}{5}$. The coin is flipped N times for some even integer N.

   (a) Using Markov's inequality, provide a bound on the probability of observing $\frac{N}{2}$ or more heads.

   (b) Using Chebyshev's inequality, provide a bound on the probability of observing $\frac{N}{2}$ or more heads. Express your answer in terms of N.

   (c) For $N \in \{2, 4, \ldots, 20\}$, in a single plot, show the bounds from part (a) and (b), as well as the *exact* probability of observing $\frac{N}{2}$ or more heads.