MA3466 Tutorial Sheet 4, outline solutions, q2[1]

22 March 2010

1. (C&T 2.16) Bottleneck. Given a Markov chain $X_1 \to X_2 \to X_3$ where the sets of outcomes have cardinalities $|\mathcal{X}_1| = n$, $|\mathcal{X}_2| = k$ and $|\mathcal{X}_3| = m$ with k is less than both m and n.

   (a) Show that the dependence of X_1 and X_3 is limited by the bottleneck by proving that $I(X_1; X_3) \le \log k$.

   (b) Evaluate $I(X_1; X_3)$ for $k = 1$ and conclude that no dependence can survive such a bottleneck.

   *Solution:*Well by the data processing inequality $I(X_1; X_3) \le I(X_1; X_2)$ but

   $$I(X_1; X_2) \le H(X_2) - H(X_2|X_1) \le H(X_2) \le \log k \tag{1}$$

   where the second inequality replaces the entropy by its maximum. Now, if $k = 1$ then $I(X_1; X_3) = 0$ since $\log 1 = 0$.

2. Consider the game of &$^{\text{TM}}$: a board game with numbered squares, the idea being to race your token along the square by flipping a coin, heads you advance one place, tails you stay put. Let Y_1, Y_2 and Y_3 be you position after one, two and three turns. Clearly $Y_1 \to Y_2 \to Y_3$, show this by calculating the conditional probabilites and showing

   $$p(x, y, z) = p(x)p(x|y)p(z|y) \tag{2}$$

   Calculate $I(Y_1; Y_2)$ and $I(Y_1; Y_3)$.

3. (C&T 2.43b) A fair six-sided dice is rolled. What is the mutual information between the top face of the dice and the side most facing you?

   *Solution:*Well

   $$I(X; Y) = H(X) - H(X|Y) \tag{3}$$

   Here we use X to denote the front of the dice and Y the front. Obviously

   $$H(X) = \log 6 = 1 + \log 3 \tag{4}$$

   since all possibilities are equally likely and six is the number of possibilities. On the other hand, if the top has a given value, say one, the front has four equally likely values, so,

   $$H(X|Y = 1) = \log 4 = 2 \tag{5}$$

   and $H(X|Y)$ is the average of the $H(X|Y = y)$ values; since they are all equally likely

   $$H(X|Y) = 2 \tag{6}$$

   so

   $$I(X; Y) = H(X) - H(X|Y) = \log 3 - 1. \tag{7}$$

4. (C&T 2.29) For random variables X, Y and Z prove the following inequalities and find conditions for equality.

   (a) $H(X, Y|Z) \ge H(X|Z)$

   (b) $I(X, Y; Z) \ge I(X; Z)$

   (c) $H(X, Y, Z) - H(X, Y) \le H(X, Z) - H(X)$

   (d) $I(X; Z|Y) \ge I(Z; Y|X) - I(Z; Y) + I(X; Z)$

   *Solution:*For $H(X, Y|Z) \ge H(X|Z)$ we have

   $$H(X, Y|Z) = H(X|Z) + H(Y|X, Z) \ge H(X|Z) \tag{8}$$

   using the chain rule. There is equality if $H(Y|X, Z) = 0$, that is, Y is a function of X and Z. Next $I(X, Y; Z) \ge I(X; Z)$; now

   $$I(X, Y; Z) = I(X; Z) + I(Y; Z|X) \ge I(X; Z) \tag{9}$$

   with equality if $I(Y; Z|X) = 0$, that is, if Y and Z are conditionally independent given X. For $H(X, Y, Z) - H(X, Y) \le H(X, Z) - H(X)$ well, using the chain rule

   $$H(X, Y, Z) - H(X, Y) = H(Z|X, Y) \tag{10}$$

   and, similarily

   $$H(X, Z) - H(X) = H(Z|X) \ge H(Z|X, Y) \tag{11}$$

   with equality when $H(Z|X) - H(Z|X, Y) = I(Z; Y|X) = 0$ which happens when Z and Y are conditionally independent given X. Finally $I(X; Z|Y) \ge I(Z; Y|X) - I(Z; Y) + I(X; Z)$; well, by the chain rule for mutual information in different orders

   $$\begin{aligned} I(X, Y; Z) &= I(X; Z|Y) + I(Z; Y) \\ I(X, Y; Z) &= I(Y; Z|X) + I(Z; X) \end{aligned} \tag{12}$$

   so $I(X; Z|Y) + I(Z; Y) = I(Y; Z|X) + I(Z; X)$ and moving the $I(Z; Y)$ over the equals we see that the inequality in the question is actually an equality.

[1] Conor Houghton, `houghton@maths.tcd.ie`, see also `http://www.maths.tcd.ie/~houghton/MA3466`