Homework 6  
due March 15, 2017 (at 18:00) [in PDF with LaTeX to Tong Wang twang92@illinois.edu]

As I announced, HWs are regarded as substitutes for Discussion, so you may submit your ‘draft’ electronically for my comments/suggestions/hints (by the time reasonably before the deadline). Also you can freely discuss with your friends as long as it facilitates your understanding. I hope all of you will get 100%.

1. Professor A is often in bad mood these days because his experiments do not go as he expects. They say he is in bad mood 75% of time when students go to see him. He has a postdoc who is in good mood 80% of the time when his boss is in good mood, but only 30% when his boss is in bad mood. 

   (1) What is the convenient information (or entropy) to measure the amount of information you can get about the mood of the professor from the mood of his postdoc? 

   (2) Is it possible to estimate this information from the data given in the problem? If so, get it in bits.

2. Demonstrate the following (so-called) chain rule:

\[ H(A \lor B \lor C) = H(A) + H(B \mid A) + H(C \mid A \lor B). \]  

Here, the probability for the joint event \( A \lor B \) is \( \{p(a_i, b_j)\} \).

3. A Markov chain is a (discrete time) stochastic process defined by the transition probability \( p_{ij} = P_{i \to j} \) that describe the transition from state \( j \) to \( i \) occurs with probability \( p_{ij} \) in, say, one second. Notice that \( p_{ij} \geq 0 \) and \( \sum_i p_{ij} = 1 \), because the state must surely go to some (other) state wherever it is at time \( t \). If \( p_i(t) \) is the probability that the state is \( i \) at time \( t \), then

\[ p_j(t + 1) = \sum_{i=1}^{n} p_{ji} p_i(t). \]  

Let us write this as \( p(t + 1) = \pi p(t) \). Assume that the Markov chain has a unique equilibrium distribution \( \{q_i\} \):

\[ q_j = \sum_i p_{ji} q_i, \]  

or \( q = \pi q \). Define the Kullback-Leibler entropy \( K \) as

\[ K(p, q) = \sum p_i \log \frac{p_i}{q_i}, \]  

with which we already encountered in Sanov’s theorem. Show the ‘second law,’

\[ K(p, q) \geq K(\pi p, q). \]  

(You must have immediately realized that this is related to Jensen. )
4 [Monty Hall Problem]
There are A, B, C three boxes of which one contains a prize. The player is asked to choose one box. Then, Monty opens one of the empty boxes. What is the amount of information (in bits) the player gets from Monty’s opening one box?