## Homework Set no. 9

1. Mutual information for correlated normals (EIT 9.3). Find the mutual information I(X;Y), where

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2 \left( 0, \begin{bmatrix} \sigma^2 & \rho \sigma^2 \\ \rho \sigma^2 & \sigma^2 \end{bmatrix} \right).$$

Evaluate I(X; Y) for  $\rho = 1, \rho = 0$ , and  $\rho = -1$ , and comment. Hint: this is a mechanical problem.

The purpose of the problem is to learn how to compute the mutual information between two continuous random variables, and to get a feel for what the information looks like for a very common case.

2. Mutual information We know that the mutual information I(X,Y) between two discrete random variables can be written as H(Y) - H(Y | X). Let the conditional distribution of Y given X, P(Y | X), be fixed, and assume that we can vary only the distribution of X, P(X). Is it true that a distribution  $P_{\max}(X)$  (not necessarily unique) that maximizes H(Y) also maximizes I(X,Y)? Prove your answer formally (that is, if your answer is "yes", you must provide a proof, while, if your answer is "no", you can provide either a proof or a counterexample.)

This is a difficult problem if you try a brute-force approach. Hint: if the answer were "yes", what would be the consequences? Where, in this course, could we apply this theorem ? Ask yourself the same questions assuming that the answer is "no".

- 3. The Channel coding Theorem At the end of the proof of the Channel Coding Theorem we used certain properties of random variables, which we did not prove in class (or maybe we did ?)
  - Show that, if the average of a set of non-negative numbers is equal to a, then at least half of the numbers must be less than or equal to 2a.
  - If X is a discrete random variable, prove that it takes a value less than or equal to E(X) with probability greater than zero, where E(X) is the expected value of X.

(Note: this is not a difficult problem, but needs a bit of care.)

4. Rate distortion function with infinite distortion. Find the rate distortion function  $R(D) = \min I(X; \hat{X})$  for  $X \sim \text{Bernoulli}(\frac{1}{2})$  and distortion

$$d(x,\hat{x}) = \begin{cases} 0, & x = \hat{x}, \\ 1, & x = 1, \hat{x} = 0, \\ \infty, & x = 0, \hat{x} = 1. \end{cases}$$