INFORMATION THEORY EXERCISE

(1) Consider the problem of minimizing $E[(Y-X)^2]$ when X has pdf

$$p(x) = \frac{1}{2\sqrt{2\pi}} \left(e^{-\frac{1}{2}x^2} + e^{-\frac{1}{2}(x-1)^2} \right)$$

and Y is constrained to have a discrete marginal distribution, concentrated on just two points (which are, however, subject to choice).

(a) Find the optimal joint distribution and the conditional distributions for $Y \mid X$ and $X \mid Y$.

(b) This problem can be thought of as equivalent to coding X into a form where it can be sent through an error-free one-bit per time unit channel (like a telegraph key). How close does your solution come to achieving the optimal 1-bit rate?

[Hint: This $p(x)$ is the convolution of a normal with a certain discrete distribution. This should allow you to exploit the optimality result from the notes.]

(2) Consider the problem of minimizing $E[(Y-X)^2]$ when X is a random variable distributed on the integers from 0 to 5 with binomial probabilities, i.e.

$$\frac{1}{32}, \frac{5}{32}, \frac{5}{16}, \frac{5}{16}, \frac{5}{32}, \frac{1}{32}.$$

We choose Y as a function of X subject only to the constraint that it can take on only 3 values, which we will call y_1, y_2, y_3. What these values are is subject to choice.

(a) What are the optimal y_i's and what is the optimal mapping from X to Y?

(b) What is the channel capacity of a device (like the one considered in this problem) that can send, without error, one of 3 values each period?

(c) What is the rate of information transmission in the solution you have proposed as the optimal mapping in this problem?

(d) If the problem were to choose the joint distribution of X and Y optimally without any constraint except that the rate of information flow had to be no greater than that of this 3-symbol channel, what would the optimal choice of the joint distribution be? (Of course here you are still taking the same distribution of X as given.) [This part of the problem may be really hard. You can give up after trying for a reasonable time and writing down any insights you have on the subject.]

[Hint: You will probably want to use the fact that it is optimal to choose $y_i = E[X \mid Y = y_i]$. You will probably need to use numerical methods. A fairly efficient numerical optimization routine, `csminwel.m`, will be on the course web site. If you can't get convergence of an optimization, shrewd guesses about the optimal joint distribution of X and Y are ok as answers. Answers will be graded according to how close they come to the best answer submitted. Finally, you might start with the simpler problem where X has a three-point distribution on $\{0, 1, 2\}$ with probabilities $\{.25, .5, .25\}$ and y has to have a two-point distribution. That might even be solvable analytically.