Problem Set 5
Due: March 22, 2006

1. Consider an exponentially distributed random variable X with parameter λ. Let F be the distribution function. Find the real number μ that satisfies: $F(\mu) = \frac{1}{2}$. This number μ is called the *median* of the random variable.

2. One of two wheels of fortune, A and B, is selected by the flip of a fair coin, and the wheel chosen is spun once to determine the experimental value of random variable X. Random variable Y, the reading obtained with wheel A, and random variable W, the reading obtained with wheel B, are described by the PDFs

$$f_Y(y) = \begin{cases} 1, & 0 < y \leq 1; \\ 0, & \text{otherwise,} \end{cases} \qquad \text{and} \qquad f_W(w) = \begin{cases} 3, & 0 < w \leq \frac{1}{3}; \\ 0, & \text{otherwise.} \end{cases}$$

   If we are told the experimental value of X was less than $\frac{1}{4}$, what is the conditional probability that wheel A was the one selected?

3. A 6.041 graduate opens a new casino in Las Vegas and decides to make the games more challenging from a probabilistic point of view. In a new version of roulette, each contestant spins the following kind of roulette wheel. The wheel has radius r and its perimeter is divided into 20 intervals, alternating red and black. The red intervals (along the perimeter) are twice the width of the black intervals (also along the perimeter). The red intervals all have the same length and the black intervals all have the same length. After the wheel is spun, the center of the ball is equally likely to settle in any position on the edge of the wheel; in other words, the angle of the final ball position (marked at the ball's center) along the wheel's perimeter is distributed uniformly between 0 and 2π radians.

   (a) What is the probability that the center of the ball settles in a red interval?

   (b) Let B denote the event that the center of the ball settles in a black interval. Find the conditional PDF $f_{Z|B}(z)$, where Z is the distance, along the perimeter of the roulette wheel, between the center of the ball and the edge of the interval immediately clockwise from the center of the ball?

   (c) What is the unconditional PDF $f_Z(z)$?

   Another attraction of the casino is the Gaussian slot machine. In this game, the machine produces independent identically distributed (IID) numbers $X_1, X_2, ...$ that have normal distribution $\mathcal{N}(0, \sigma^2)$. For every i, when the number X_i is positive, the player receives from the casino a sum of money equal to X_i. When X_i is negative, the player pays the casino a sum of money equal to $|X_i|$.

   (d) What is the standard deviation of the net total gain of a player after n plays of the Gaussian slot machine?

   (e) What is the probability that the absolute value of the net total gain after n plays is greater than $2\sqrt{n}\sigma$?

4. Let a continuous random variable X be uniformly distributed over the interval $[-1, 1]$. Derive the PDF $f_Y(y)$ where:

(a) $Y = \sin(\frac{\pi}{2}X)$

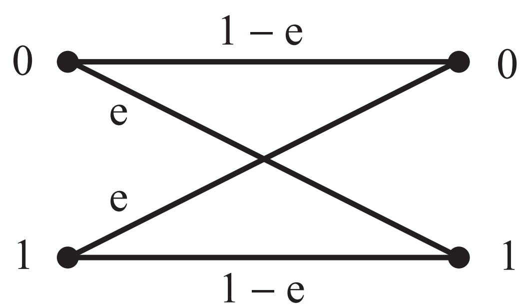
(b) $Y = \sin(2\pi X)$

5. **A Communication Example** During lecture, we looked at a simple point-to-point communication example. The task was to transmit messages from one point to another, by using a noisy channel. We modeled the problem probabilistically as follows:

   - A binary message source generates independent successive messages $M_1, M_2, \cdots$: each message is a discrete random variable that takes the value 1 with probability p and the value 0 with probability $1-p$. For a single message, we write the PMF as:

$$p_M(m) = \begin{cases} 1-p & \text{If } m = 0 \\ p & \text{If } m = 1 \end{cases}$$

   - A binary symmetric channel acts on an input bit to produce an output bit, by flipping the input with "crossover" probability e, or transmitting it correctly with probability $1-e$.



For a single transmission, this channel can be modeled using a conditional PMF $p_{Y|X}(y|x)$, where X and Y are random variables for the channel input and output bits respectively. We then have:

$$p_{Y|X}(y|x) = \begin{cases} 1-e & \text{If } y = x. \\ e & \text{If } y \neq x. \end{cases}$$

Next, we make a series of assumptions that make the above single-transmission description enough to describe multiple transmissions as well. We say that transmissions are:

(a) Independent: Outputs are *conditionally* independent from each other, *given the inputs.*

(b) Memoryless: Only the current input affects the current output.

(c) Time-invariant: We always have the same conditional PMF.

We write:

$$\begin{aligned} &p_{Y_1,Y_2,\cdots|X_1,X_2,\cdots}(y_1, y_2, \cdots | x_1, x_2, \cdots) \\ &\overset{(a)}{=} p_{Y_1|X_1,X_2,\cdots}(y_1|x_1, x_2, \cdots) \cdot p_{Y_2|X_1,X_2,\cdots}(y_2|x_1, x_2, \cdots) \cdots \\ &\overset{(b)}{=} p_{Y_1|X_1}(y_1|x_1) \cdot p_{Y_2|X_2}(y_2|x_2) \cdots \\ &\overset{(c)}{=} p_{Y|X}(y_1|x_1) \cdot p_{Y|X}(y_2|x_2) \cdots \end{aligned}$$

Any transmission scheme through this channel must transform messages into channel inputs (encoding), and transform channel outputs into estimates of the transmitted messages (decoding). For this problem, we will encode each message separately (more elaborate schemes look at blocks and trails of messages), and generate a sequence of n channel input bits. The encoder is therefore a map:

$$\begin{array}{rcl} \{0,1\} & \longmapsto & \{0,1\}^n \\ M & \longrightarrow & X_1, X_2, \cdots, X_n \end{array}$$

The decoder is simply a reverse map:

$$\begin{array}{rcl} \{0,1\}^n & \longmapsto & \{0,1\} \\ Y_1, Y_2, \cdots, Y_n & \longrightarrow & \hat{M} \end{array}$$

Note that we use the "hat" notation to indicate an estimated quantity, but bare in mind that M and $\hat{M}$ are two distinct random variables. The complete communication problem looks as follows:

Finally, to measure the performance of any transmission scheme, we look at the probability of error, i.e. the event that the estimated message is different than the transmitted message:

$$\mathbf{P}(\text{error}) = \mathbf{P}(\hat{M} \neq M).$$

(a) **No encoding:**
The simplest encoder sends each message bit directly through the channel, i.e. $X_1 = X = M$. Then, a reasonable decoder is to use the output channel directly as message estimate: $\hat{M} = Y_1 = Y$. What is the probability of error in this case?

(b) **Repetition code with majority decoding rule:**
The next thing we attempt to do is to send each message $n > 1$ times through this channel. On the decoder end, we do what seems natural: decide 0 when there are more 0s than 1s, and decide 1 otherwise. This is a "majority" decoding rule, and we can write it as follows (making the dependence of $\hat{M}$ on the channel outputs explicit):

$$\hat{M}(y_1, \cdots, y_n) = \begin{cases} 0 & \text{If } y_1 + \cdots + y_n < n/2. \\ 1 & \text{If } y_1 + \cdots + y_n \geq n/2. \end{cases}$$

Analytical results:

i. Find an expression of the probability of error as a function of e, p and n. [Hint: First use the total probability rule to divide the problem into solving $\mathbf{P}(\text{error}|M = 0)$ and $\mathbf{P}(\text{error}|M = 1)$, as we did in the lecture.]

ii. Choose $p = 0.5$, $e = 0.3$ and use your favorite computational method to make a plot of $\mathbf{P}(\text{error})$ versus n, for $n = 1, \cdots, 15$.

iii. Is majority decoding rule optimal (lowest $\mathbf{P}(\text{error})$) for all p? [Hint: What is the best decoding rule if $p = 0$?].

Simulation:

i. Set $p = 0.5$, $e = 0.3$, generate a message of length 20.

ii. Encode your message with $n = 3$.

iii. Transmit the message, decode it, and write down the value of the message error rate (the ratio of bits in error, over the total number of message bits).

iv. Repeat Step 3 serveral times, and average out all the message error rates that you obtain. How does this compare to your analytical expression of $\mathbf{P}(\text{error})$ for this value of n?

v. Repeat Steps 3 and 4 for $n = 5, 10, 15$, and compare the respective average message error rates to the corresponding analytical $\mathbf{P}(\text{error})$.

(c) **Repetition code with maximum a posteriori (MAP) rule:**

In Part b, majority decoding was chosen almost arbitrarily, and we have alluded to the fact that it might not be the best choice in some cases. Here, we claim that the probability of error is in fact minimized when the decoding rule is as follows:

$$\hat{M}(y_1, \cdots, y_n) = \begin{cases} 0 & \text{If } \mathbf{P}(M = 0 | Y_1 = y_1, \cdots, Y_n = y_n) \\ & \qquad > \mathbf{P}(M = 1 | Y_1 = y_1, \cdots, Y_n = y_n) \\ 1 & \text{Otherwise.} \end{cases}$$

This decoding rule is called "maximum a posterior" or MAP, because it chooses the value of M which maximizes the posterior probability of M given the channel ouput bits $Y_1 = y_1, \cdots, Y_n = y_n$ (another term for Bayes' rule).

i. Denote by N_0 the number of 0s in $y_1, \cdots, y_n$, and by N_1 the number of 1s. Express the MAP rule as an inequality in terms of N_0 and N_1, and as a function of e and p. [Hint: Use Bayes' rule to decompose the posterior probability. Note that $p_{Y_1, \cdots, Y_n}(y_1, \cdots, y_n)$ is a constant during one decoding.]

ii. Show that the MAP rule reduces to the majority rule if $p = 0.5$.

iii. Give an interpretation of how, for $p \neq 0.5$, the MAP rule deviates from the majority rule. [Hint: Use the simplification steps of your previous answer, but keep p arbitrary.]

iv. (Optional) Prove that the MAP rule minimizes $\mathbf{P}(\text{error})$.

G1†. Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of *independent* and *identically distributed* random variables. We define random variables $S_1, S_2, \ldots, S_n, \ldots$ $(n = 1, 2 \ldots)$ in the following manner: $S_n = X_1 + X_2 + \cdots + X_n$. Find

$$E[X_1 \mid S_n = s_n, S_{n+1} = s_{n+1}, \ldots]$$