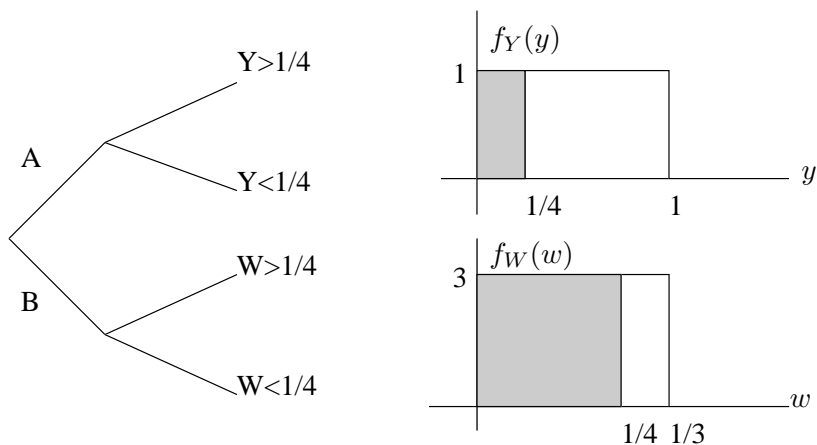**Problem Set 5 Solutions**
Due: March 22, 2006

1. We are given that $F(\mu) = \frac{1}{2}$

$$\Rightarrow \quad \int_0^\mu \lambda e^{-\lambda x} dx = \frac{1}{2}$$

$$\Rightarrow \quad \int_0^\mu \lambda e^{-\lambda x} dx = -e^{-\lambda x}|_0^\mu = 1 - e^{-\lambda \mu} = \frac{1}{2}$$

$$\Rightarrow \quad -\lambda \mu = \ln \frac{1}{2}$$

$$\Rightarrow \quad \mu = \frac{\ln 2}{\lambda}$$

2. First we can draw a tree with the the following branches:

Then, using the PDFs given in the question we can determine two probabilities that are clearly relevant for this question and give branch labels for the tree:

$$\mathbf{P}\left(Y \leq \frac{1}{4}\right) = \int_0^{\frac{1}{4}} f_Y(y)\, dy = \frac{1}{4}$$

$$\mathbf{P}\left(W \leq \frac{1}{4}\right) = \int_0^{\frac{1}{4}} f_W(w)\, dw = \frac{3}{4}$$

Finally, Bayes's Rule gives

$$\mathbf{P}\left(A \mid X < \frac{1}{4}\right) = \frac{\mathbf{P}(A)\mathbf{P}(X \leq \frac{1}{4} \mid A)}{\mathbf{P}(X \leq \frac{1}{4})} = \frac{\mathbf{P}(A)\mathbf{P}(Y \leq \frac{1}{4})}{\mathbf{P}(X \leq \frac{1}{4})}$$

$$= \frac{\mathbf{P}(A)\mathbf{P}(Y \leq \frac{1}{4})}{\mathbf{P}(A)\mathbf{P}(X \leq \frac{1}{4} \mid A) + \mathbf{P}(B)\mathbf{P}(X \leq \frac{1}{4} \mid B)}$$

$$= \frac{\mathbf{P}(A)\mathbf{P}(Y \leq \frac{1}{4})}{\mathbf{P}(A)\mathbf{P}(Y \leq \frac{1}{4}) + \mathbf{P}(B)\mathbf{P}(W \leq \frac{1}{4})}$$

$$= \frac{\frac{1}{2} \cdot \frac{1}{4}}{\frac{1}{2} \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{3}{4}} = \frac{1}{4}.$$

3. (a) We know that the total length of the edge for red interval is two times that for black inteval. Since the ball is equally likely to fall in any position of the edge, probability of falling in a red interval is $\frac{2}{3}$.

   (b) Conditioned on the ball having fallen in a black interval, the ball is equally likely to fall anywhere in the interval. Thus, the PDF is

   $$f_{Z|\text{black interval}}(z) = \begin{cases} \frac{15}{\pi r} & , \quad z \in [0, \frac{\pi r}{15}] \\ 0 & , \quad \text{otherwise} \end{cases}$$

   (c) Since the ball is equally likely to fall on any point of the edge, we can see it is twice as likely for $z \in [0, \frac{\pi r}{15})$ than $z \in \left[\frac{\pi r}{15}, \frac{2\pi r}{15}\right]$. Therefore, intuitively, let

   $$f_Z(z) = \begin{cases} 2h & , \quad z \in [0, \frac{\pi r}{15}) \\ h & , \quad z \in \left[\frac{\pi r}{15}, \frac{2\pi r}{15}\right] \\ 0 & , \quad \text{otherwise} \end{cases}$$

   Using the fact that $\int_{-\infty}^{\infty} f_Z(z) = 1$,

   $$(2h)(\frac{\pi r}{15}) + (h)(\frac{\pi r}{15}) = 1 \Rightarrow h = \frac{5}{\pi r}$$

   $$f_Z(z) = \begin{cases} \frac{10}{\pi r} & , \quad z \in [0, \frac{\pi r}{15}) \\ \frac{5}{\pi r} & , \quad z \in \left[\frac{\pi r}{15}, \frac{2\pi r}{15}\right] \\ 0 & , \quad \text{otherwise} \end{cases}$$

   (d) The total gains (or losses), T, equals to the sum of all X_i, i.e. $T = X_1 + X_2 + \cdots + X_n$, Since all the X_i's are independent of each other, and they have the same Gaussian distribution, the sum will also be a Gaussian with

   $$\mathbf{E}[T] = \mathbf{E}[X_1] + \mathbf{E}[X_2] + \cdots + \mathbf{E}[X_n] = 0$$

   $$\text{var}(T) = \text{var}(X_1) + \text{var}(X_2) + \cdots + \text{var}(X_n) = n\sigma^2$$

   Therefore, the standard deviation for T is $\sqrt{n}\sigma$.

   (e)

   $$\begin{aligned} \mathbf{P}(|T| > 2\sqrt{n}\sigma) &= \mathbf{P}(T > 2\sqrt{n}\sigma) + \mathbf{P}(T < -2\sqrt{n}\sigma) \\ &= 2\mathbf{P}(T > 2\sqrt{n}\sigma) \\ &= 2\left(1 - \Phi\left(\frac{2\sqrt{n}\sigma - \mathbf{E}[T]}{\sigma_T}\right)\right) \\ &= 2(1 - \Phi(2)) \simeq 0.0454. \end{aligned}$$

4. (a) $Y = g(X) = \sin(\frac{\pi}{2}X)$. Because $g(x)$ is a monotonic function for $-1 < x < 1$, we can define an inverse function $h(y) = \frac{2 \arcsin y}{\pi}$ and use the PDF formula given in lecture:

   $$f_Y(y) = f_X(h(y))\left|\frac{dh(y)}{dy}\right| \quad \text{for } -1 < y < 1$$

$$
\begin{aligned}
&= f_X\left(\frac{2\arcsin y}{\pi}\right)\left|\frac{2}{\pi\sqrt{1-y^2}}\right| \\
&= \frac{1}{2}\cdot\frac{2}{\pi\sqrt{1-y^2}} \\
&= \frac{1}{\pi\sqrt{1-y^2}}
\end{aligned}
$$

The final answer is:

$$
f_Y(y) = \begin{cases} 0 & y < -1 \\ \frac{1}{\pi\sqrt{1-y^2}} & -1 \le y < 1 \\ 0 & 1 \le y \end{cases}
$$

(b) $Y = \sin(2\pi X)$. In rearranging the equation, we find $X = \frac{\arcsin Y}{2\pi}$. This is a many-to-one mapping of X to Y. That is, given a value of Y, it is not possible to pinpoint the corresponding value of X. For example, if $Y = 1$, X could have been $\frac{1}{4}$ or $\frac{-3}{4}$. This means we cannot use the formula used in part (a).

The CDF is still the means to achieving the PDF, but we take a graphical approach instead. First, let's consider the extreme cases. By virtue of the $\sin X$ function, the value of Y varies from -1 to 1. It is obvious that no value of the random variable Y can be less than -1, so $F_Y(y \le -1) = 0$. Also, every value of Y is less than 1, so $F_Y(y \ge 1) = 1$.

For $-1 \le y \le 0$, consider the following diagram:
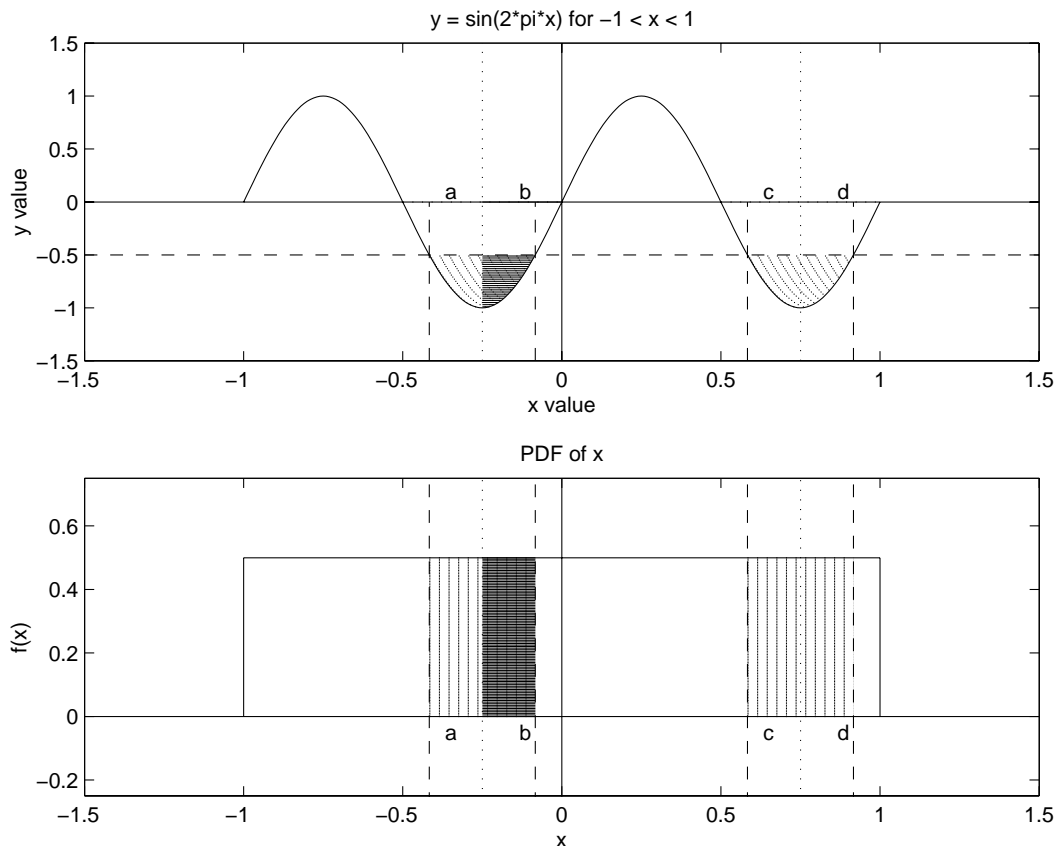
The value b indicates the conventional response of $\arcsin y$ which is between $[-\frac{\pi}{2}, \frac{\pi}{2}]$. In other words, by rearranging the original equation:

$$b = \frac{\arcsin y}{2\pi}$$
$$-\frac{1}{4} \leq b \leq \frac{1}{4}$$

So, we now have $F_Y(y) = \mathbf{P}((a \leq X \leq b) \cup (c \leq X \leq d))$, where the event on the right is the shaded region in the above PDF of X. By symmetry, and mutual exclusivity, this shaded region can be expressed as four times the more darkly shaded region of the PDF.

$$
\begin{aligned}
F_Y(y| -1 \leq y \leq 0) &= \mathbf{P}((a \leq X \leq b) \cup (c \leq X \leq d)) \\
&= 4\mathbf{P}\left(-\frac{1}{4} \leq X \leq b\right) \\
&= 4(0.5)\left[b - \left(-\frac{1}{4}\right)\right] \\
&= \frac{\arcsin y}{\pi} + \frac{1}{2}.
\end{aligned}
$$

For $0 \leq y \leq 1$, consider the following diagram:
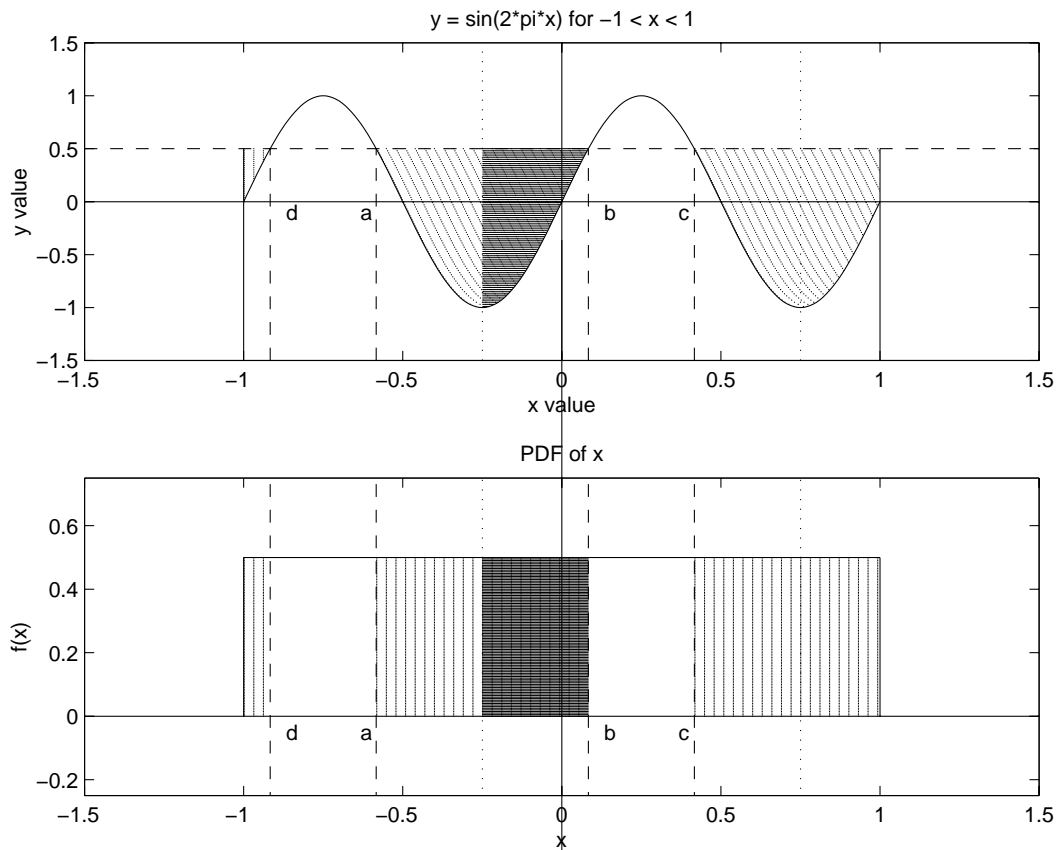
By an analogous argument to the previous case, we find that $F_Y(y|0 \leq y \leq 1)$ is represented by the shaded region from -1 to d, a to b, and c to 1. Once again, however, this is exactly

four times the more darkly shaded region from $-\frac{1}{2}$ to b. So, the expression for the CDF is the same as in the case $-1 \leq y \leq 0$.

$$F_Y(y|0 \leq y \leq 1) = \frac{\arcsin y}{\pi} + \frac{1}{2}$$

To summarize:

$$F_Y(y) = \begin{cases} 0 & y \leq -1 \\ \frac{\arcsin y}{\pi} + \frac{1}{2} & -1 \leq y < 1 \\ 1 & 1 \leq y \end{cases}$$

Now use the identity $\frac{d}{dy} \arcsin y = \frac{1}{\sqrt{1-y^2}}$ and differentiate with respect to y:

$$f_y(y) = \frac{d}{dy} F_Y(y) = \begin{cases} 0 & y < -1 \\ \frac{1}{\pi\sqrt{1-y^2}} & -1 \leq y < 1 \\ 0 & 1 \leq y \end{cases}$$

5. *Analytical Results*

   (a) **No encoding:**

$$\begin{aligned} \mathbf{P}(\text{error}) &= \mathbf{P}(\hat{M} = 1, M = 0) + \mathbf{P}(\hat{M} = 1, M = 0) \\ &= \mathbf{P}(Y = 1, X = 0) + \mathbf{P}(Y = 0, X = 1) \\ &= \mathbf{P}(Y = 1|X = 0)\mathbf{P}(X = 0) + \mathbf{P}(Y = 0|X = 1)\mathbf{P}(X = 1) \\ &= e(1 - p) + ep \end{aligned}$$

   (b) **Repetition encoding:**

   i. Similar to the previous case, the errors can be attributed to the conditions $\hat{M} = 1, M = 0$ and $\hat{M} = 0, M = 1$. Assuming we encode each bit by repeating it $n = 2m + 1, m = 0, 1..$ times, we get

$$\begin{aligned} \mathbf{P}(\text{error}) &= \mathbf{P}(\hat{M} = 1, M = 0) + \mathbf{P}(\hat{M} = 1, M = 0) \\ &= \mathbf{P}(Y_1 + Y_2 \ldots Y_n \geq m + 1|M = 0)(1 - p) + \mathbf{P}(Y_1 + Y_2 \ldots Y_n \leq m|M = 1)p \\ &= \mathbf{P}(\text{atleast } (m + 1) \text{ errors}|M = 0)(1 - p) + \mathbf{P}(\text{atleast } (m + 1) \text{ errors}|M = 0)p \\ &= \sum_{k=m+1}^{n} \binom{n}{k} e^k (1 - e)^{(n-k)} \end{aligned}$$

   Note: It can be seen that the probability of error is independent of the priori probabilities p and $(1 - p)$. The error varies only as a function of n

   ii. The variation of the error probability as a function of n is shown in Figure 1

   iii. Even though, the error is independent of p, in degenerate cases such as $p = 1, p = 0$, the majority decoding does not give us any advantage. It is easier to always output $\hat{M} = 1, \hat{M} = 0$ respectively.

   Simulations The matlab code for simulating the communication channel is attached with the solutions. Figure 2 shows the close agreement between the analytical results (indicated by $\circ$) and the simulated error rates (indicated by $*$).
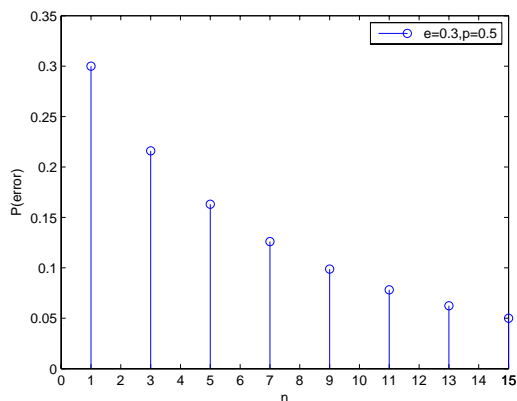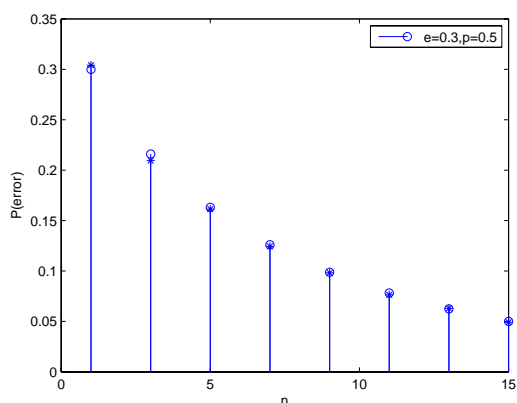
Figure 1: Analytical Results

Figure 2: Simulation results

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

i. Using Baye's rule to compute the posterior probabilities, we get

$$
\begin{aligned}
\mathbf{P}(M = 0 | Y_1 = y_1, \cdots, Y_n = y_n) &= \frac{\mathbf{P}(Y_1 = y_1, \cdots, Y_n = y_n | M = 0)\mathbf{P}(M = 0)}{\mathbf{P}(Y_1 = y_1, \cdots, Y_n = y_n)} \\
&= e^{N_1}(1 - e)^{N_0}(1 - p) \\
\text{Similarly,} \mathbf{P}(M = 1 | Y_1 = y_1, \cdots, Y_n = y_n) &= \frac{\mathbf{P}(Y_1 = y_1, \cdots, Y_n = y_n | M = 1)\mathbf{P}(M = 1)}{\mathbf{P}(Y_1 = y_1, \cdots, Y_n = y_n)} \\
&= e^{N_0}(1 - e)^{N_1}(p)
\end{aligned}
$$

Since, the denominators are the same, the MAP decision rule reduces to

$$
\hat{M}(y_1, \cdots, y_n) = \begin{cases} 0 & \text{If } e^{N_1}(1 - e)^{N_0}(1 - p) \geq e^{N_0}(1 - e)^{N_1}(p) \\ 1 & \text{Otherwise.} \end{cases}
$$

ii. When, $p = 0.5$, the decision rule reduces to

$$\hat{M}(y_1, \cdots, y_n) = \begin{cases} 0 & \text{If } e^{N_1}(1-e)^{N_0} \geq e^{N_0}(1-e)^{N_1} \\ 1 & \text{Otherwise.} \end{cases}$$

The test expression can be rewritten as $(1-e)^{N_0 - N_1} \geq e^{N_0 - N_1}$ If we assume that $e \leq 0.5$, then $(1-e) \geq e$. Therefore, we can re-write the test as,

$$\hat{M}(y_1, \cdots, y_n) = \begin{cases} 0 & \text{If } N_0 > N_1 \\ 1 & \text{Otherwise.} \end{cases}$$

iii. In this case, we see that the decision rule and hence the probability of error depends oupon the *a-priori* probabilities p and $1-p$. In particular, if we consider the degenerate cases again $p = 0, p = 1$, we see from the rule that we always decide in favor of $\hat{M} = 0, \hat{M} = 1$ irrespective of the recieved bits. This is in contrast to the majority decoding, that still has to count the number of $1s$ in the output.

iv. Here we resort to a more intuitive proof of the statement that also illustrates the concept of 'risk' in decision making (A mathematically rigorous proof may be found in 6.011 and 6.432). Consider the case when we have no received data and make the decision based entirely on prior probabilities $\mathbf{P}(M = 0) = 1 - p$ and $\mathbf{P}(M = 1) = p$. If we decide $\hat{M} = 1$, then we have the $1-p$ probability of being wrong. Similarily if we chose $\hat{M} = 0$, we have a probability p of being wrong. We choose $\hat{M}$ to minimize the risk of being wrong or maximize the probability of being right. Thus we choose

$$\hat{M} = \arg \max_{M=(0,1)} \mathbf{P}(X = M)$$

When we are given the data Y, we have to deal with the *a-posteriori* probabilities $\mathbf{P}(M = 1|Y), \mathbf{P}(M = 0|Y)$ instead of *a-priori* probabilities $\mathbf{P}(M = 0), \mathbf{P}(M = 1)$ and the argument remains unchanged. Thus, to minimize probability or being wrong, or maximize the probability of being right, we choose

$$\hat{M} = \arg \max_{M=(0,1)} \mathbf{P}(X = M|Y)$$

G1[†]. Note that we can rewrite $E[X_1 \mid S_n = s_n, S_{n+1} = s_{n+1}, \ldots]$ as follows:

$$
\begin{aligned}
&E[X_1 \mid S_n = s_n, S_{n+1} = s_{n+1}, \ldots] \\
=\ & E[X_1 \mid S_n = s_n, X_{n+1} = s_{n+1} - s_n, X_{n+2} = s_{n+2} - s_{n+1}, \ldots] \\
=\ & E[X_1 \mid S_n = s_n],
\end{aligned}
$$

where the last equality holds due to the fact that the X_i's are independent.

We also note that

$$E[X_1 + \cdots + X_n \mid S_n = s_n] = E[S_n \mid S_n = s_n] = s_n$$

It follows from the linearity of expectations that

$$E[X_1 + \cdots + X_n \mid S_n = s_n] = E[X_1 \mid S_n = s_n] + \cdots + E[X_n \mid S_n = s_n]$$

Because the X_i's are identically distributed, we have the following relationship:

$$E[X_i \mid S_n = s_n] = E[X_j \mid S_n = s_n], \text{ for any } 1 \le i \le n, 1 \le j \le n.$$

Therefore,

$$E[X_1 + \cdots + X_n \mid S_n = s_n] = nE[X_1 \mid S_n = s_n] = s_n \quad \Rightarrow \quad E[X_1 \mid S_n = s_n] = \frac{s_n}{n}.$$