Problem 1

Optimal Code Lengths that Require One Bit above Entropy

Let $X$ be a binary chance variable with probability mass function

$$P_X(x) = \begin{cases} 1 - \delta & \text{for } x = x_0, \\ \delta & \text{for } x = x_1, \end{cases}$$

with $0 < \delta < 1$. Then an optimal code assigns to both symbols a codeword of length 1, e.g.,

$$c(x_0) = 0, \quad c(x_1) = 1,$$

where $c(x)$ denotes the codeword for $x$. The expected length of this code is

$$L = 1.$$

However, the entropy of $X$ is $H(X) = H_b(\delta)$ where $H_b(\cdot)$ denotes the binary entropy function. Since $\inf_\delta H_b(\delta) = 0$ it follows that for any $\epsilon > 0$ there exists a $\delta \in (0,1)$ such that $H_b(\delta) < \epsilon$. Thus

$$L = 1 > 1 + H_b(\delta) - \epsilon = H(X) + 1 - \epsilon.$$

Problem 2

Shannon Code

a) Let $c_i$ be the codeword corresponding to $F_i$. In order to show that the code is prefix-free, we first note that since $p_1 \geq p_2 \geq \cdots \geq p_m$, we have $l_1 \leq l_2 \leq \cdots \leq l_m$. We further note that for any $j > i,$

$$F_j - F_i = \sum_{k=i}^{j-1} p_k \geq \sum_{k=i}^{j-1} 2^{-l_k},$$

where the inequality follows because $l_k = \lceil \log \frac{1}{p_k} \rceil \geq \log \frac{1}{p_k}.$

In the following we show that for $j > i$ and thus $l_i \leq l_j$ the codeword $c_i$ cannot be a prefix of $c_j$. To this end, note that $c_i$ can only be a prefix of $c_j$, if the first $l_i$ symbols of $c_j$ are identical to $c_i$, or equivalently if the first $l_i$ bits in the binary representation of $F_j$ are the same as the first $l_i$ bits in the binary representation of $F_i$. But this can only be the case if $F_j - F_i < 2^{-l_i}$ which by (1) is not possible, and consequently, $c_i$ can never be a prefix of $c_j$. This concludes the proof.

Now we look at the expected length:
\[
\log \frac{1}{p_i} \leq l_i < \log \frac{1}{p_i} + 1,
\]
\[
p_i \log \frac{1}{p_i} \leq p_i l_i < p_i \log \frac{1}{p_i} + p_i,
\]
\[
\sum_{i=1}^{m} p_i \log \frac{1}{p_i} \leq \sum_{i=1}^{m} p_i l_i < \sum_{i=1}^{m} p_i \log \frac{1}{p_i} + \sum_{i=1}^{m} p_i,
\]
\[
H(X) \leq L < H(X) + 1.
\]

b)

\[
F_1 = 0 \equiv 0.0000 \quad l_1 = 1 \implies 0,
\]
\[
F_2 = 0.5 \equiv 0.1000 \quad l_2 = 2 \implies 10,
\]
\[
F_3 = 0.75 \equiv 0.1100 \quad l_3 = 3 \implies 110,
\]
\[
F_4 = 0.875 \equiv 0.1110 \quad l_4 = 3 \implies 111.
\]

The codewords are then \{0, 10, 110, 111\}.