Problem 1

a) [A, 35%], [B, 20%], [C, 15%], [D, 15%], [E, 8%], [F, 7%] $n = 6$

Huffman ([A, 35%], [B, 20%], [C, 15%], [D, 15%], [E, 8%], [F, 7%])

2 smallest f's: $f_6 = 7\%$, $f_5 = 8\%$

Huffman ([F, 7+8%], [A, 35%], [B, 20%], [C, 15%], [D, 15%])

2 smallest f's: $f_1 = 15\%$, $f_4 = 15\%$

Huffman ([F, 30%], [A, 35%], [B, 20%], [D, 15%])

2 smallest f's: $f_4 = 15\%$, $f_3 = 20\%$

Huffman ([D, 35%], [F, 30%], [A, 35%])

2 smallest f's: $f_2 = 30\%$, $f_1 = 35\%$

Huffman ([F, 65%], [A, 35%])

2 smallest f's: $f_2 = 35\%$, $f_1 = 65\%$

Huffman ([A, 100%])

code[A] = ""

code[F] = 0

code[A] = 1

code[D] = code[F] + 0 = 00

code[F] = code[F] + 1 = 01

code[B] = code[D] + 0 = 000

code[D] = code[D] + 1 = 001

code[C] = code[F] + 0 = 010

code[F] = code[F] + 1 = 011

code[E] = code[F] + 0 = 0110

code[F] = code[F] + 1 = 0111

$f_1 = 15\%$, $f_3 = 15\%$

[F, 30%], [A, 35%], [B, 20%], [C, 15%]

$f_4 = 15\%$, $f_3 = 20\%$

[C, 35%], [F, 30%], [A, 35%]

$f_2 = 30\%$, $f_3 = 35\%$

[F, 65%], [C, 35%]

$f_2 = 35\%$, $f_1 = 65\%$

[C, 100%]

code[C] = ""

code[F] = 0

code[C] = 1

code[A] = code[F] + 0 = 00

code[F] = code[F] + 1 = 01

code[B] = code[C] + 0 = 10

code[C] = code[C] + 1 = 11

code[D] = code[F] + 0 = 010

code[F] = code[F] + 1 = 011

code[F] = code[F] + 0 = 0110

code[F] = code[F] + 1 = 0111

Code #1:

A	1	D	001
B	000	E	0110
C	010	F	0111

Code #2:

A	00	D	010
B	10	E	0110
C	11	F	0111

b) For code #1: $1(0.35) + 3(0.2) + 3(0.15) + 3(0.15) + 4(0.08) + 4(0.07)$

$= 2.45$

For code #2: $2(0.35) + 2(0.2) + 2(0.15) + 3(0.15) + 4(0.08) + 4(0.07)$

$= 2.45$

c) No, there does not exist a prefix-free code with a smaller accepted codeword length. By Lemma 1, an optimal prefix-free code will have $f_x > f_y$ and $\text{length}(c_x) \leq \text{length}(c_y)$ for any x, y. By Lemma 2, E and F (the letters with the smallest frequencies) must have equal length, and the smallest length it could be in this case while remaining prefix-free is 4. A, B, C, and D must be less than 4, so they must have lengths 1, 3, 3, 3 or lengths 2, 2, 2, 3. This is because at least one variation of each length codeword can't stand on its own to allow the other codewords to build off of it without causing duplicate prefixes. Length 1 has 2 variations, and length 2 has 4 variations.