ECE 515 FALL 2026 ASSIGNMENT 3 SOLUTIONS

1. a) Interaction Information

$p(x=0) = p(x=1) = 1/2$

$p(y=0) = p(y=1) = 1/2$

$p(z=0) = p(z=1) = 1/2$

$H(X) = H(Y) = H(Z) = 1$ bit

$p(0,0) = \frac{8}{20}$ $p(0,1) = \frac{2}{20}$ $p(1,0) = \frac{2}{20}$ $p(1,1) = \frac{8}{20}$

$H(XY) = H(YZ) = H(XZ) = 1.722$ bits

$H(XYZ) = -\frac{14}{20}\log\frac{2}{20} - \frac{6}{20}\log\frac{1}{20} = 2.357$ bits

$I(X;Y;Z) = H(X) + H(Y) + H(Z) - H(XY) - H(YZ)$

$- H(YZ) + H(XYZ)$

$= 3 - 3(1.7222) + 2.357 = 0.191$ bit

the sign is positive so this is redundancy not synergy

this comes from the fact that 70% of the mass is on the two outcomes where all three agree

b) there is shared information among the three assets so the portfolio has less diversification than having three individual assets suggests

$I(X;Y) = .278$ bit and $I(X;Y|Z) = .087$ bit

so Z explains much of the information between X and Y

2(a) $D[p(x,y,z) \| p(x)p(y)p(z)]$

$$= \sum_x \sum_y \sum_z p(x_i, y_j, z_k) \log \left[\frac{p(x_i, y_j, z_k)}{p(x_i)p(y_j)p(z_k)} \right]$$

$$= -H(XYZ)$$

$$- \sum_x \sum_y \sum_z p(x_i, y_j, z_k) \log p(x_i) p(y_j) p(z_k)$$

using the identity $\log(abc) = \log a + \log b + \log c$

$$D[p(x,y,z) \| p(x)p(y)p(z)]$$

$$= -H(XYZ) + H(X) + H(Y) + H(Z)$$

(b) $D[p(x,y,z) \| p(x)p(y)p(z)] = 0$ iff

$p(x,y,z) = p(x)p(y)p(z)$ for all x, y, z

i.e. if X and Y and Z are independent

3(a) $p(x_1)=.5 \quad p(x_2)=.25 \quad p(x_3)=.125 \quad p(x_4)=.125$

$q(x_1)=.5 \quad q(x_2)=a \quad q(x_3)=c \quad q(x_4)=d$

$$H(p,q) = H(p(X)) + D[p(X)\|q(X)]$$

Since $D[p(X)\|q(X)] \geq 0$

$H(p,q)$ is minimized when

$H(p,q) = H(p(X))$

which occurs when $q(X)=p(X)$

∴ choose $a=.25 \quad b=.125 \quad c=.125$

then $$H(p,q) = -\sum_{i=1}^{4} p(x_i)\log q(x_i)$$

$$= -.5\log.5 - .25\log.25 - .25\log.125$$

$$= .5 + .5 + .75 = 1.75 \text{ bits}$$

Since $D[p(X)\|q(X)] = 0$

$$H(p,q) = H(p(X)) + D[p(X)\|q(X)]$$

(b) $c = .25$ so $q(x_4) = .25$

$$H(p,q) = -\sum_{i=1}^{4} p(x_i) \log q(x_i)$$

$$= -.5 \log .5 - .25 \log a - .125 \log b - .125 \log .25$$

$$= .75 - .25 \log a - .125 \log b$$

$$\sum_{i=1}^{4} q(x_i) = 1$$

$$\therefore\ a + b = .25 \quad \text{or} \quad b = .25 - a$$

$$H(p,q) = .75 - .25 \log a - .125 \log (.25 - a)$$

to minimize the cross entropy, set

$$\frac{dH(p,q)}{da} = 0$$

$$\frac{dH(p,q)}{da} = -.25\left(\frac{1}{a \ln 2}\right) - .125 \frac{1}{(.25-a)\ln 2}(-1)$$

$$\frac{-.25}{a \ln 2} + \frac{.125}{(.25-a)\ln 2} = 0$$

$$.25(.25-a) = .125a$$

$$.0625 - .25a = .125a$$

$$a = \frac{.0625}{.375} = .1667$$

so $b = .0833$

$$H(p,q) = .75 - .25\log .1667 - .125\log .0833$$

$$= .75 + .6462 + .4481$$

$$= 1.844 \text{ bits}$$

$$D[p(x)\|q(x)] = \sum_{i=1}^{4} p(x_i)\log\frac{p(x_i)}{q(x_i)}$$

$$= .5\log\frac{.5}{.5} + .25\log\frac{.25}{.1667} + .125\log\frac{.125}{.0833} + .125\log\frac{.125}{.25}$$

$$= 0 + .25\log 1.5 + .125\log 1.5 + .125\log 0.5$$

$$= .375(.585) - .125 = .0944 \text{ bit}$$

from (a) $H(p(x)) = 1.75$ bits

so $H(p(x)) + D[p(x)\|q(x)) = 1.844$ bits $= H(p,q)$

(c) . In (a) $H(p,q) = 1.75$ bits and in (b)

$H(p,q) = 1.844$ bits which is .0944 bit larger

The cross entropy is a measure of how different two probability distributions are

Since $H(p,q)$ in (a) is smaller than in (b)

$q(x)$ in (a) is close to $p(x)$ than

$q(x)$ in (b)

4 Determine which of the following codes is uniquely decodable and which are prefix codes

a) $\{0, 10, 11\}$ prefix code so UD

b) $\{0, 01, 11\}$ UD but not prefix

c) $\{0, 01, 10\}$ NOT UD so not prefix

010 could be 0, 10 or 01, 0

d) $\{0, 01\}$ UD but not prefix

e) $\{110, 11, 10, 01\}$ Not UD so not prefix

110110 could be 110, 110 or 11, 01, 10

f) $\{00, 01, 10, 11\}$ prefix code so UD

(g) $\{0, 00, 000, 0000\}$ NOT UD so not prefix

5. Let X be a random variable with N equiprobable outcomes

a) optimal binary prefix code

there will exist an optimal prefix code such that the longest and shortest codewords differ by at most 1

Let $\ell_s = \lfloor \log_2 N \rfloor$ and $\ell_l = \lceil \log_2 N \rceil$

and $d = N - 2^{\lfloor \log_2 N \rfloor}$

then the optimal code has

$2d$ codewords of length ℓ_l

and

$N - 2d$ codewords of length ℓ_s

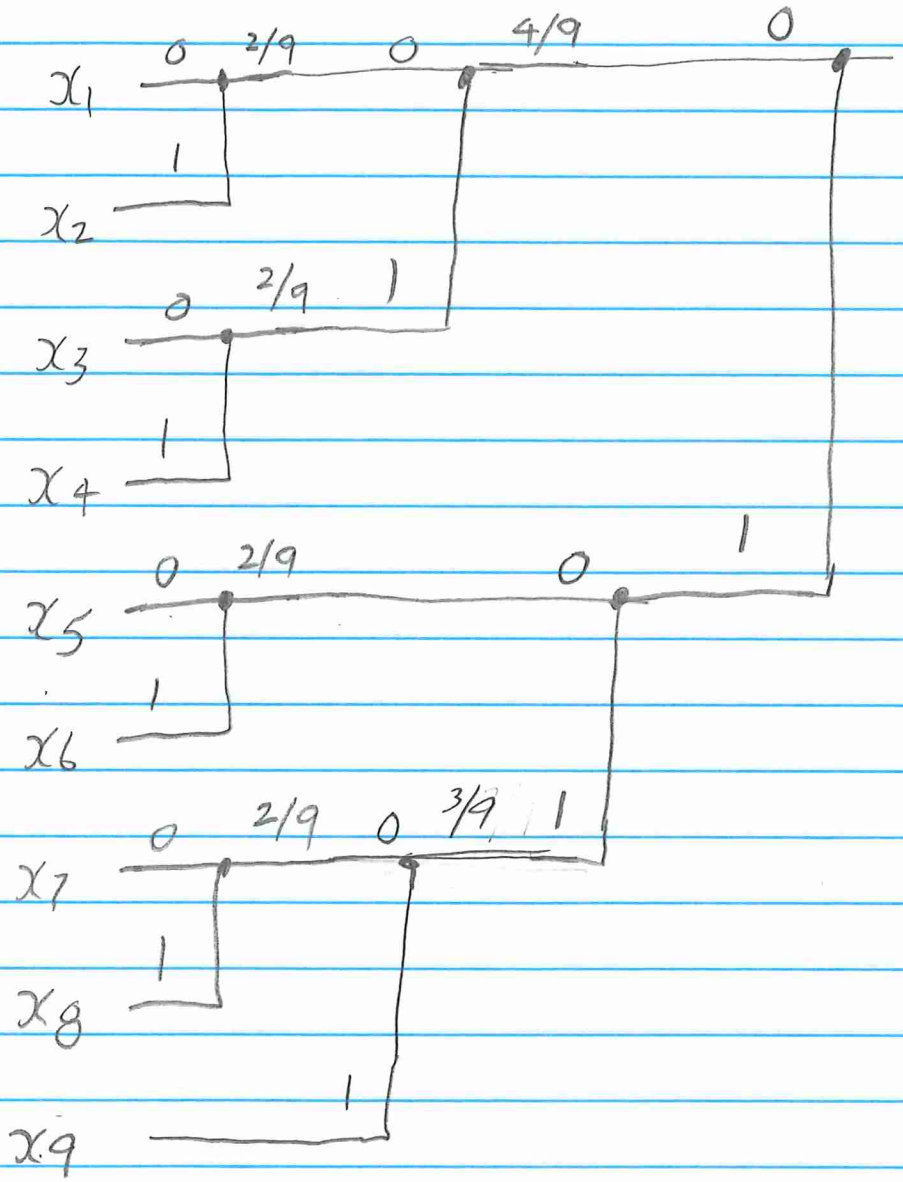
$$L(C) = \frac{1}{N}\left(2d\lceil \log_2 N \rceil + (N-2d)\lfloor \log_2 N \rfloor\right)$$

$$= \frac{1}{N}\left(N\lfloor \log_2 N \rfloor + 2d\right)$$

$$= \lfloor \log_2 N \rfloor + \frac{2d}{N}$$

b) $N = 9$ $\quad l_s = \lfloor \log_2 9 \rfloor = 3$ $\quad l_l = \lceil \log_2 9 \rceil = 4$

$d = 9 - 2^3 = 1$



x_1	000
x_2	001
x_3	010
x_4	011
x_5	100
x_6	101
x_7	1100
x_8	1101
x_9	111

c) when N is a power of 2

$$l_s = l_f = \log_2 N$$

$$d = N - 2^{\log_2 N} = 0$$

$\therefore$ all codewords have length $\log_2 N$

$$H(X) = -\sum_{i=1}^{N} \frac{1}{N} \log_2 \frac{1}{N} = \log_2 N$$

$$= L(C)$$

6. This is a Huffman code for an infinite alphabet

Considering the subset of symbols

$$\{x_j : j > i\}$$

the cumulative probability is

$$\sum_{j>i} 0.9\,(0.1)^{j-1} = 0.9\,(0.1)^{i-1}\left(\frac{1}{1-0.1}\right)$$

$$= (0.1)^{i-1}$$

since this is less then

$$p(x_i) = 0.9\,(0.1)^{i-1}$$

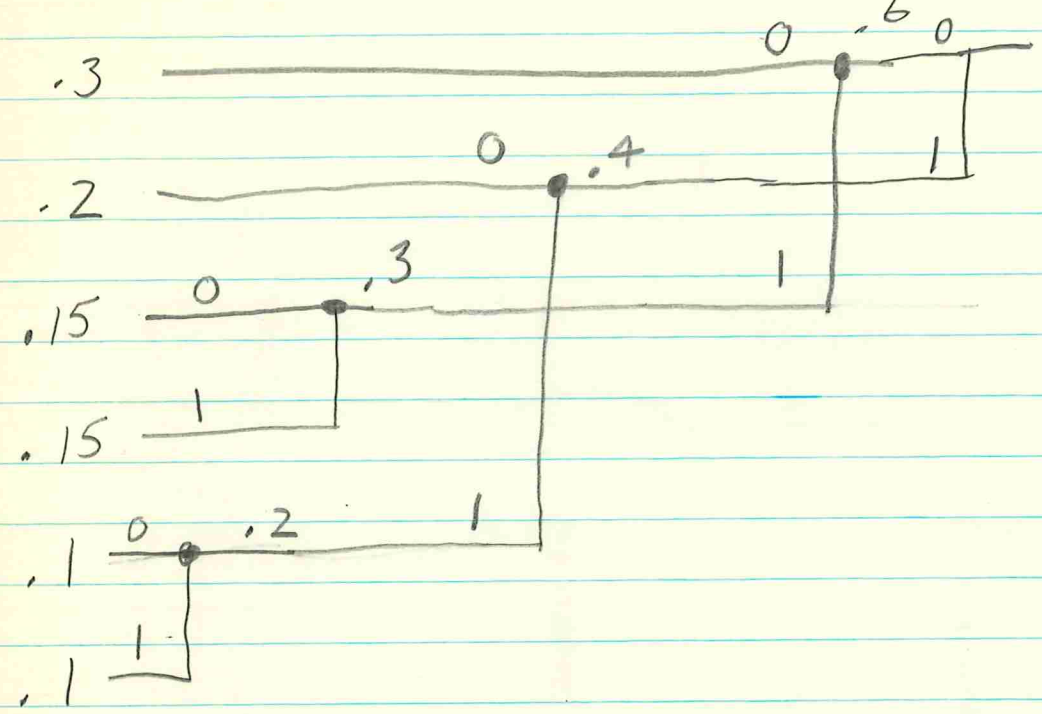
the Huffman algorithm will always merge the last two terms so the codeword lengths are

1, 2, 3, 4, 5, ...

One possible code is

0, 10, 110, 1110, 11110, ...

7. first code - binary Huffman code



C_1 00

C_2 10

C_3 010

C_4 011

C_5 110

C_6 111

$L(C) = .5(2) + .5(3) = 2.5$ bits

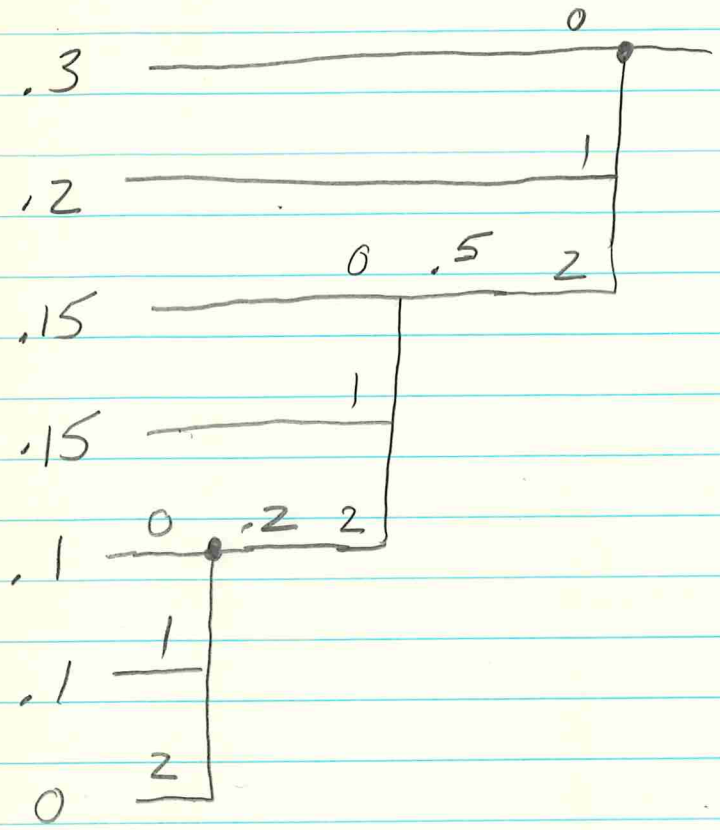
$H(X) = .521 + .464 + .821 + .664 = 2.47$ bits

$\zeta = 98.8\%$

ternary Huffman code

needs $J + C(J-1)$ symbols

$3 + 2(2) = 7$ so extend the source by 1 symbol



C_1 0

C_2 1

C_3 20

C_4 21

C_5 221

C_6 220

$L(C) = .5(1) + .3(2) + .2(3) = 1.7$ trits

or $(\log_2 3)1.7 = 2.69$ bits

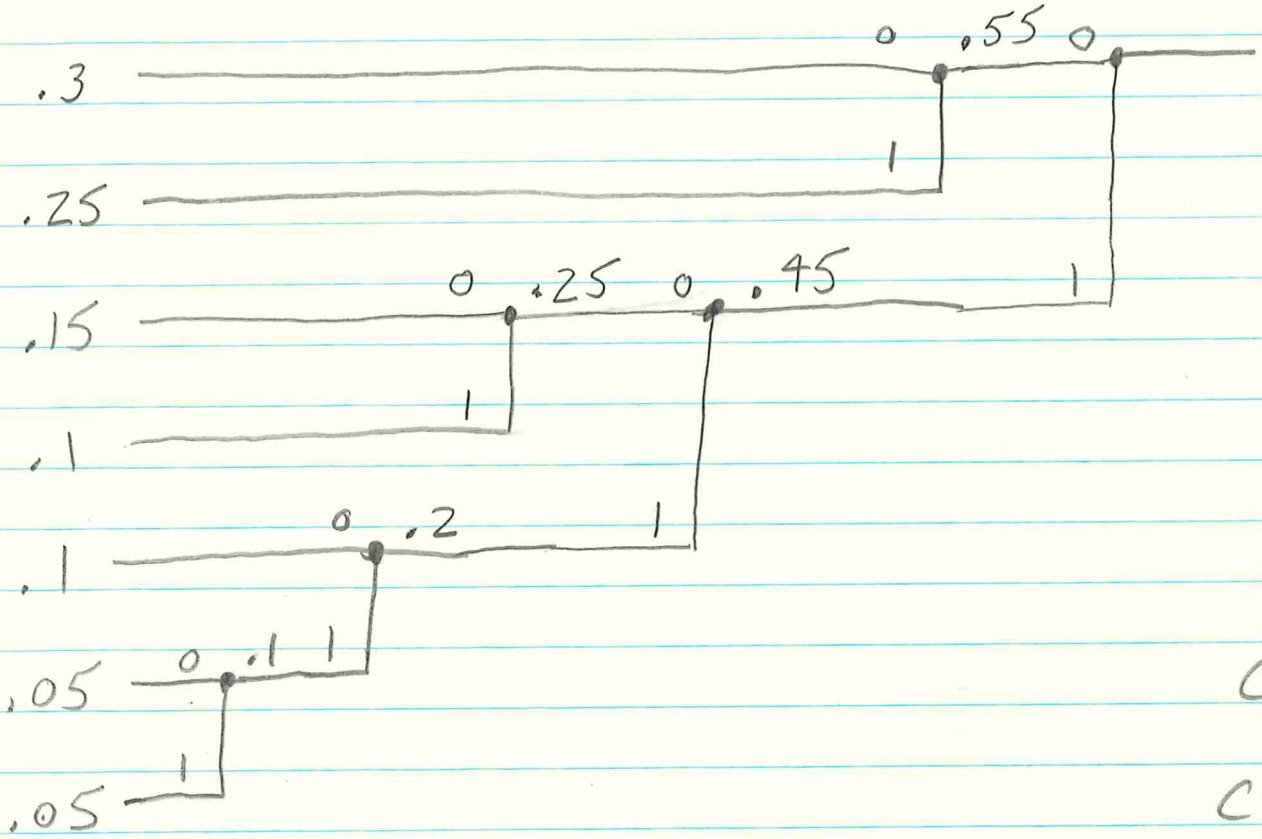
$H(X) = .329 + .293 + .518 + .419 = 1.56$ trits

or $(\log_2 3)\,1.56 = 2.47$ bits

$\eta = 91.8\%$

Thus the binary code is better.

second code- binary Huffman code



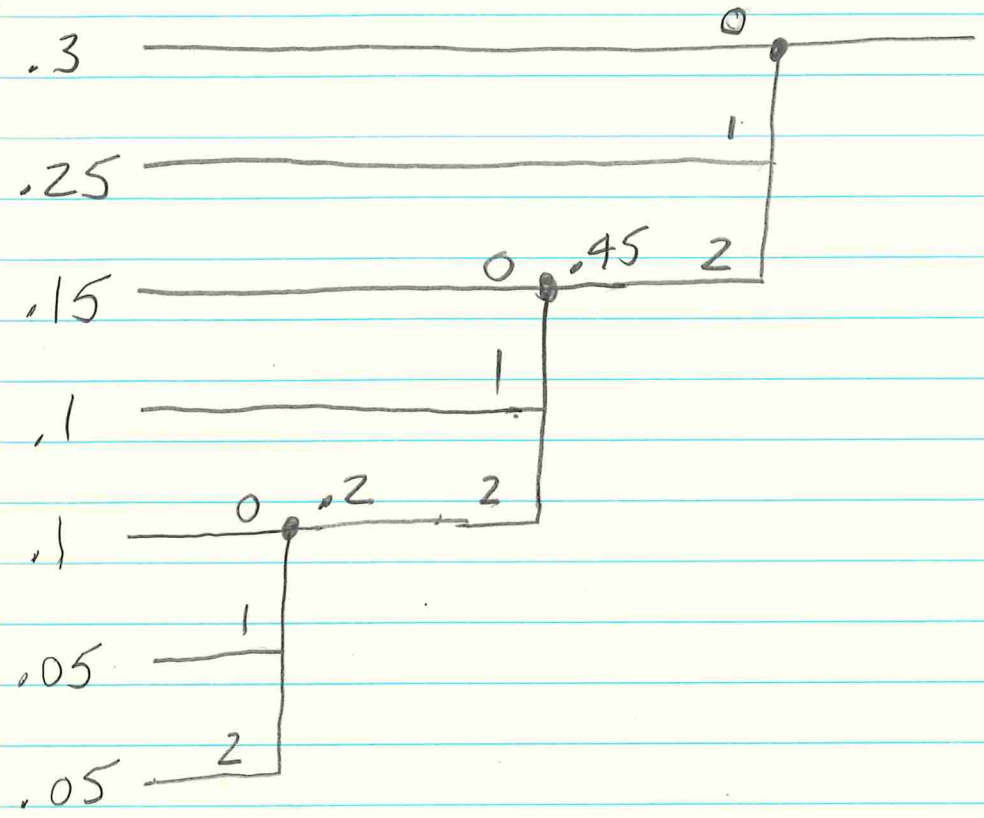
$L(C) = .55(2) + .35(3) + .1(4) = 2.55$ bits

$H(X) = .521 + .5 + .411 + .664 + .432 = 2.53$ bits

$\eta = 99.1\%$

ternary Huffman code

$3 + 2(2) = 7$ so no extra symbols are required



c_1 0

c_2 1

c_3 20

c_4 21

c_5 220

c_6 221

c_7 222

$L(C) = .55(1) + .25(2) + .2(3) = 1.65$ trits

or $(\log_2 3)1.65 = 2.62$ bits

$H(X) = .329 + .315 + .259 + .419 + .273 = 1.595$ trits

$\} = 96.6\%$

The binary code is again better than the ternary code

Length Uniqueness

Source 1: binary - yes

ternary - no

alternative code: {0, 10, 11, 12, 20, 21}

Note that $L(C) = .3(1) + .7(2) = 1.7$ trits

so both codes are optimal

Source 2: binary - yes, all possible choices lead to the same lengths

ternary - yes