Question 1

Question 1(a)

1) $p(X \geq a) \leq \frac{E[X]}{a}$ given $E[X] = 2000$, $a = 2400$

$p(X \geq 2400) \leq \frac{2000}{2400}$ or $\frac{5}{6}$ #

2) $P(|X-\mu| \geq k \times \sigma) \leq \frac{1}{k^2}$ given $\mu = 2000$, $\sigma = 200$ Hence, $k = \frac{2400-2000}{200} = 2$

$p(1600 \leq X \leq 2400) \geq 1 - \frac{1}{2^2}$ or $\frac{3}{4}$ #

Question 1(b)

1) $p(X = head) < \frac{1}{8}$, coin flipped N times $N \in N$

$p(X^N \geq \frac{N}{4}) \leq \frac{E[X^N]}{N/4} = \frac{1}{2}$ #

2) $p < \frac{1}{8}$, coin flipped N times

$\mu = \frac{N}{8}$, $\sigma^2 = \frac{7N}{64}$ $k = \frac{\frac{N}{4} - \frac{N}{8}}{\frac{\sqrt{7N}}{8}} = \frac{N}{\sqrt{7N}}$

$p(X^N \geq \frac{N}{4}) \leq \frac{7}{N}$

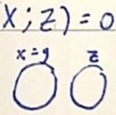
3) plot

Question 2 : Markov Chain

Question 2(a)

1) yes it's possible, for example let X be a random variable where $x \in \{0,1\}$ with $P(X=0) = P(X=1) = \frac{1}{2}$ then let $Y = X$ and $Z = X$. Hence $I(X;Y) = I(X;Z) = H(X)$ since Y and Z are complety determined by X

2) yes. let X be random v. that $P(X=0) = P(X=1) = \frac{1}{2}$ and $Y = X$. Z is a variable that independent to Y. Hence, $I(X;Y) = H(X) = H(Y)$ and $I(X;Z) = 0$, then $I(X;Y) \geq I(X;Z)$



3) according to chain rule for mutual information :

$I(X;Y,Z) = I(X;Y) + I(X;Z|Y) = I(X;Z) + I(X;Y|Z)$

Hence $I(X;Y) + I(X;Z|Y) = I(X;Z) + I(X;Y|Z)$

according to markov chain assumption where $I(X;Z|Y) = 0$

Proof : $I(X;Z|Y) = H(Z|Y) - H(Z)$

$= E\left[\log_2 \frac{p(X,Z|Y)}{p(X|Y)\,p(Z|Y)}\right] = E[\log_2 1] = 0$

Then $I(X;Y) = I(X;Z) + I(X;Y|Z)$ Hence $I(X;Y) \geq I(X;Z)$

4) according to the chain rule on previous question, we get

$I(X;Y) + I(X;Z|Y) = I(X;Z) + I(X;Y|Z)$

$I(X;Y|Z) = I(X;Y) - I(X;Z)$ since $I(X;Z)$ must be equal or more than 0 Hence $I(X;Y) \geq I(X;Y|Z)$

this is because when we know Z, the uncertainty can only decreze or stay the same. In other word, the dependence between X and Y can't be increased by the observation of a downstream variable.

Question 2(b)

1) we know that $I(X;Y,Z) = H(X) - H(X|Y,Z)$ and $I(X;T) = H(X) - H(X|T)$ since $p(t|y,z,x) = p(t|y,z)$, according Markov, this means $H(X|Y,Z) \leq H(X|T)$, then we substitute and get $I(X;Y,Z) = H(X) - H(X|Y,Z) \geq H(X) - H(X|T) = I(X;T)$ #

2) In order for $I(X;Y,Z) = I(X;T)$ knowing T must give same amount of information with Y,Z or $H(X|Y,Z) = H(X|T)$. Hence, T must be a function of Y, Z & that $p(t|y,z) = p(t|y,z,x)$

SUBJECT : NO : DATE : / /

Question 2 (c)

1) By definition of mutual information, we get

$$I(X_1; X_2, \dots, X_n) = H(X_1) - H(X_1 | X_2, \dots, X_n)$$

where $H(X_1 | X_2, \dots, X_n) = H(X_1 | X_2, \dots, X_{n-1}) + H(X_1 | X_2, \dots, X_{n-2}, X_n) + \dots + H(X_1 | X_2, X_n) + H(X_1 | X_n)$

and we plug it back, we get: $I(X_1; X_2, \dots, X_n) = H(X_1) - H(X_1 | X_2, \dots, X_{n-1}) - H(X_1 | X_2, \dots, X_{n-2}, X_n) - \dots - H(X_1 | X_2, X_n) - H(X_1 | X_n)$

2) we can simplify expression above and get:

$$I(X_1; X_2, \dots, X_n) = H(X_1) - H(X_1 | X_n)$$

Question 3 : AEP

a) $H(X) = 0.8 \times \log_2\left(\frac{1}{0.8}\right) + 0.2 \times \log_2\left(\frac{1}{0.2}\right) = 0.7219$ bits

b) since there are two outcome and $N \in \{0, 1, 2, \dots, N\}$ the size of $A_X^N = 2^N$

c) since $|A_{X^4}| = 16$, then $H_0(X^4) = \log_2 |A_{X^4}| = 4$

d) $H(X^N) = N \sum_x p(x) \log p(x) = N H(X)$

e) As we increase N from small to large, the slope of the line become more flat meaning that it become ~~the~~ less sensitive to change in error. When N get large, such sequence occupy most of the probability mass, and are equally likely.

Question 5 : AEP

collection
- Alina (7)
- Beyonce (12)
- Cecilia ~~8~~(15)
- Derek (14)

(4)

1) a. $\log_2(4) = 2$ bits, $\{00, 01, 10, 11\}$

b. $\lceil \log_2(7) \rceil = 3$ bits, $\{000, 001, 010, 011, 100, 101, 110\}$

c. $\lceil \log_2(48) \rceil = 6$ bits

2) $H_0(X) = \lceil \log_2(48) \rceil = 6$ bits

3) a. Ensemble $A = (x, A_X, P_X)$ where $A_X = \{$Alina, Beyonce, Cecilia, Derek$\}$ and $P_X = (7/48, 12/48, 15/48, 14/48)$ such that the first element of P_X is Alina and so on.

b. $\log_2(256) = 8$ bit

c. smallest is 0

d. largest is 0.3958

4) a. $H(A) = -\left(\frac{7}{48}\log_2\frac{7}{48} + \frac{12}{48}\log_2\frac{12}{48} + \frac{15}{48}\log_2\frac{15}{48} + \frac{14}{48}\log_2\frac{14}{48}\right)$

$= 1.95$

b. $|T_N^\delta| \approx 2^{100(1.98+0.1)} \approx 2^{100 \times 2.08}$

c. no, it's not possible because the minimum of bit of uniform code with zero loss is 1.95 which is the minimum for uniform code and 1.5 bit is less than 1.95