1. a) X can take $1 \to \infty$

$$\left.\begin{aligned} P(X=1) &= \frac{1}{2} \\ P(X=2) &= \frac{1}{2}\cdot\frac{1}{2} \\ P(X=k) &= \left(\frac{1}{2}\right)^k \end{aligned}\right\} H(X) = -\sum_{n=1}^{\infty} P(n)\log_2 P(n)$$

$$H(X) = -\sum_{n=1}^{\infty}\left(\frac{1}{2}\right)^n \log_2\left(\frac{1}{2}\right)^n$$

$$= -\sum_{n=1}^{\infty}\left(\frac{1}{2}\right)^n\cdot(-n)$$

$$= \frac{\frac{1}{2}}{\left(1-\frac{1}{2}\right)^2} = 2 \text{ bits}$$

b) question sequence : most likely $\Rightarrow$ unlikely

Δ ask questions that have probability $\to \frac{1}{2}$

reduce uncertainty most efficiently

greedy algorithm

eg. 1. 1st ?

2. 2nd ?

⋮

$$E(k) = \sum_{k=1}^{\infty} P(n)\cdot n$$

$$= \sum_{k=1}^{\infty}\left(\frac{1}{2}\right)^n\cdot n = \frac{\frac{1}{2}}{\left(1-\frac{1}{2}\right)^2} = 2$$

2. r.v. X, $y = g(X)$

$$H(X, g(X)) = H(X) + \underbrace{H(g(X) \mid X)}_{} \quad {=0: X \to g(X)}$$

$$= \underbrace{H(g(X))}_{\geq 0} + \underbrace{H(X \mid g(X))}_{\geq 0}$$

$\triangle$ for discrete r.v., entropy is positive:

$\Rightarrow \boxed{H(X) \geq H(g(X))}$

$\triangle$ equality when $g(X)$ is a 1-1 function of X.

a) $y = x^2$ $\quad H(y) < H(X)$ depends on distribution.

b) $y = x^3$ $\quad H(y) = H(X)$

3. vector:

$$H(\vec{P}) = -\sum_{i=1}^{n} P(n) \log_2 P(n)$$

min: 0

max: $$-\sum_{i=1}^{N} \frac{1}{N} \log_2 \frac{1}{N} = \log_2 N$$

$$\frac{1}{N} \quad \frac{1}{N} \quad \cdots \cdots$$

$$\square \ \square \ \square \ \square \ \square \ \square \ \square \ \square \ \square \ \square$$

$$H'(P) = -\log P - P\cdot\frac{1}{P}\cdot\log e + \log(1-P) + (1-P)\frac{1}{1-P}\log e$$
$$= -\log P + \log(1-P)$$

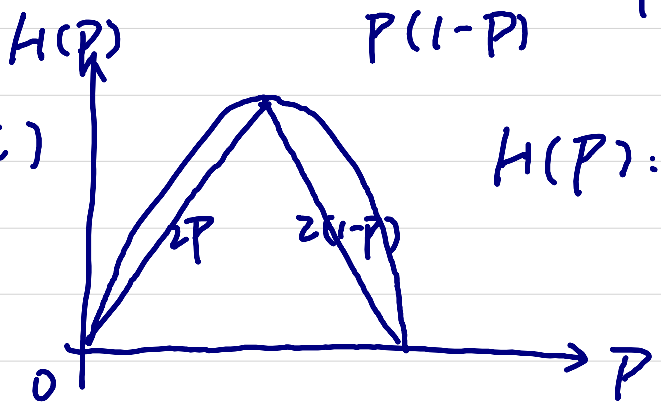
4. a) $H(P) = -P\log P - (1-P)\log(1-P)$

$$H'(P) = -\log P + \log(1-P) = \log\frac{1-P}{P}$$

b) $H''(P) = -\frac{1}{P}\log e - \frac{1}{1-P}\log e$

$$= \frac{-(1-P)-P}{P(1-P)}\log e = \frac{-\log e}{P(1-P)}$$

c)



$H(P)$: concave

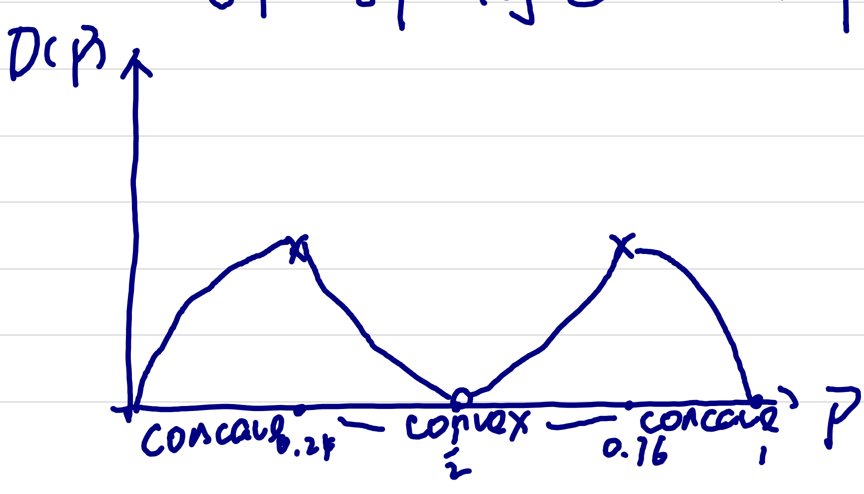
d) $D(P) = H(P) - 1 + 4(P-\frac{1}{2})^2$

prove $D(P)$ is positive in $P\in[0,1]$

$D'(P) = \log\frac{1-P}{P} + 8(P-\frac{1}{2})$ $\quad D'(P)=0 \Rightarrow P=\frac{1}{2}$ (min)

$D''(P) = \frac{-\log e}{P(1-P)} + 8 = 0$

$-8P^2 + 8P - \log e = 0 \Rightarrow P = 0.24, 0.76.$



e) write as difference.
show bounds by concavity or convexity.

$$D(p) = 1 - 2\log e \cdot (p - \tfrac{1}{2})^2 - H(p)$$

$D(p)$ is symmetrical over $p = \frac{1}{2}$. consider $p \in [0, \frac{1}{2}]$

$$D(0) = 1 - \frac{\log e}{2}, \quad D(\tfrac{1}{2}) = 0.$$

$$D'(p) = -4\log e \cdot (p - \tfrac{1}{2}) - \log \frac{1-p}{p}$$

$$D'(0) = -\infty. \quad D'(\tfrac{1}{2}) = 0.$$

$$D''(p) = -4\log e + \frac{\log e}{p(1-p)}$$

$$D''(p) = -4p + 4p^2 + 1 = 0$$

$$\Rightarrow p = \tfrac{1}{2}$$

5.

$$H(X, g(X)) = H(X) + \overbrace{H(g(X)|X)}^{=0}$$

$$= H(g(X)) + H(X|g(X))$$

$$\therefore H(X) = H(g(X)) + H(X|g(X))$$

$$\geq H(g(X))$$

6. $$H(Y|X) = \sum_{x \in \mathcal{X}} H(Y|X=x) \underbrace{P_X(x)}_{>0} = 0$$

$\because H(Y|X) = 0$

$\therefore H(Y|X=x) = 0$ for all $x \in \mathcal{X}$

$\therefore x \mapsto y$ is one-to-one mapping.

7.

$x \to f(\cdot) \to z$

A: 0000 1111 $\to$ ignore

B: 1010 0101 $\to$ 0, 1 $\quad P(z=0) = \dfrac{p^2(1-p)^2}{p^2(1-p)^2 + p^2(1-p)^2} = \dfrac{1}{2}$

C: 0001 0011 0111 $\to$ 0, 0

0010 0110 1110 $\to$ 0, 1

0100 1100 1101 $\to$ 1, 0

1000 1001 1011 $\to$ 1, 1

$$P(z_{1:2} = 00) = \frac{(1-p)^3p + (1-p)^2p^2 + (1-p)p^3}{4[(1-p)^3p + (1-p)^2p^2 + (1-p)p^3]} = \frac{1}{4}$$

$$E[K] = \sum_{k=0}^{2} k P(k)$$

$$= 0 \cdot P(K=0) + 1 \cdot \overbrace{P(K=1)}^{p^2(1-p)^2 \times 2} + 2 \cdot \overbrace{P(K=2)}^{2\times 4[(1-p)^3 p + (1-p)^2 p^2 + (1-p)p^3]}$$

$$nH(P) \overset{(a)}{=} H(X_{1:n}) \overset{(b)}{\geq} H(Z_{1:K}, K) \overset{(c)}{=} H(K) + H(Z_{1:K} | K)$$

$$\overset{(d)}{=} H(K) + E[K] \geq E[K]$$

a) $H(X_{1:n}) = n \cdot H(X) = n \cdot H(P)$ because $X_{1:n}$ is i.i.d.

b) $H(X) \geq H(f(x))$ because $(Z_{1:K}, K)$ is a func. of X

c) joint entropy

d) $H(Z_{1:K} | K) = \sum_{i=0}^{2} P_K(i) H(Z_{1:K} | K=i)$

$$= P_K(0) \cdot \underbrace{H(Z_{1:0})}_{0} + P_K(1) H(Z_{1:1}) + P_K(2) H(Z_{1:2})$$

$$= P_K(1) \cdot H(Z) + P_K(2) \cdot 2H(Z)$$

$\because X \sim B(0.1)$ $\therefore H(Z) = 1$

$\therefore H(Z_{1:K} | K) = P_K(1) \cdot 1 + P_K(2) \cdot 2 = E[K]$

e) $H(K) \geq 0$.

8. a) $\overset{H(Y|Z)-H(Y|X,Z)}{\overset{\shortparallel}{I(X;Y|Z)}} < \overset{H(Y)-H(Y|X)}{\overset{\shortparallel}{I(X;Y)}}$

$$X \longrightarrow Y = g(X) \longrightarrow Z$$

$$H(g(X)|Z) - H(g(X)|X,Z)$$

$$\le H(g(X)) - \underbrace{H(g(X)|X,Z)}_{H(g(X)|X)} = I(X;g(X))$$

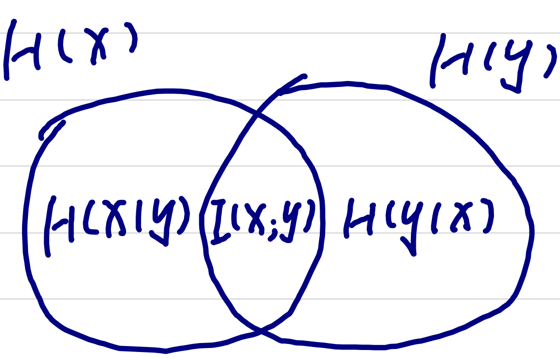
b) $I(X;Y|Z) > I(X;Y)$

$$\begin{tikzcd} & Z \arrow[d] & \\ X \arrow[r] & \oplus \arrow[r] & Y \end{tikzcd}$$

and $X \perp Z$

$$H(\overset{X+Z}{\overset{\shortparallel}{Y}}|Z) - \underbrace{H(Y|Z,X)}_{=0}$$

$$= H(X) \ge H(X) - H(X|Y) = I(X;Y)$$



9. $I(X;Y;Z) = I(X;Y) - I(X;Y|Z)$

a) $I(X;Y;Z) = H(X) - \overbrace{H(X|Y)}^{H(X,Y) - H(Y)}$

$\quad - [\underbrace{H(X|Z)}_{H(X,Z) - H(Z)} - \underbrace{H(X|Y,Z)}_{H(X,Y,Z) - H(Y,Z)}]$

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

$\quad + H(X,Y,Z) - H(Y,Z)$

b) see Q8 b).

$I(X;Y) < I(X;Y|Z)$

10. show $\log_e x \geq 1 - \frac{1}{x}$ for $x > 0$

define $f(x) = \log_e x - 1 + \frac{1}{x}$, $x > 0$

$f'(x) = \frac{1}{x} - \frac{1}{x^2} = 0 \Rightarrow x = 1$

$f''(x) = -\frac{1}{x^2} + \frac{2}{x^3} \Rightarrow f''(1) = -1 + 2 > 0$

$\therefore f(x)$ has min value at $f(1) = 0$

$\therefore \log_e x \geq 1 - \frac{1}{x}$

11.

$y \backslash x$	0	1
0	0	$\frac{1}{3}$
1	$\frac{1}{3}$	$\frac{1}{3}$

$\rightarrow P(y=0) = \frac{1}{3}$

$\rightarrow P(y=1) = \frac{2}{3}$

$P(x=0) = \frac{1}{3}$ $\quad P(x=1) = \frac{2}{3}$

$$H(x) = H(y) = -\frac{1}{3}\log\frac{1}{3} - \frac{2}{3}\log\frac{2}{3}$$

$$H(x|y) = -\sum_{x,y} P(x,y)\log P(x|y)$$

$$= -P(0,0)\log P(0,0) - P(0,1)\log P(0|1)$$

$$-P(1,0)\log P(1|0) - P(1,1)\log P(1|1)$$

$$= -\frac{1}{3}\log\frac{1}{2} - \frac{1}{3}\log 1 - \frac{1}{3}\log\frac{1}{2}$$

$H(y|x) = H(x|y)$: symmetrical

$$H(x,y) = H(x) + H(y|x)$$

$$I(x,y) = H(x) - H(x|y)$$

12.

$X \to Y \to Z$ (Markov chain)

data processing equality:

$$I(X;Y) \geq I(X;Z)$$

$$\therefore I(X;Z) = H(Y) - H(Y|X) \overset{(a)}{\leq} H(Y) \overset{(b)}{\leq} \log(k)$$

(a) $H(Y)$ and $H(Y|X) \geq 0$

(b) max entropy by uniform distribution

$$\boxed{\begin{array}{l} k=1 \Rightarrow \text{no uncertainty} \\ \qquad\quad z \text{ is a function of } x \end{array}}$$

$I(X;Y) = 0 \Rightarrow$ the degree of uncertainty are the same

$\Rightarrow$ one-to-one mapping

$\Rightarrow$ y is a function of x

.3.

a) $H(X,Y|Z) = H(X|Z) + H(Y|X,Z)$

$\geq H(X|Z)$

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

$\geq I(X;Z)$

c) $\underbrace{H(X,Y,Z) - H(X,Y)}_{H(Z|X,Y)} \leq \underbrace{H(X,Z) - H(X)}_{H(Z|X)}$

d) $I(X;Z|Y) + I(Z;Y) \geq I(Z;Y|X) + I(X;Z)$

$I(X,Y;Z) \quad = \quad I(X,Y;Z)$