

bibhasitkgr21@gmail.com

cc: Rohit

Lec - 5

$x \uparrow$
 $y \downarrow$
 x, y

is Range span of X .

$\boxed{X} : S_X \rightarrow \{H, T\}$

$\{5, 2\}$

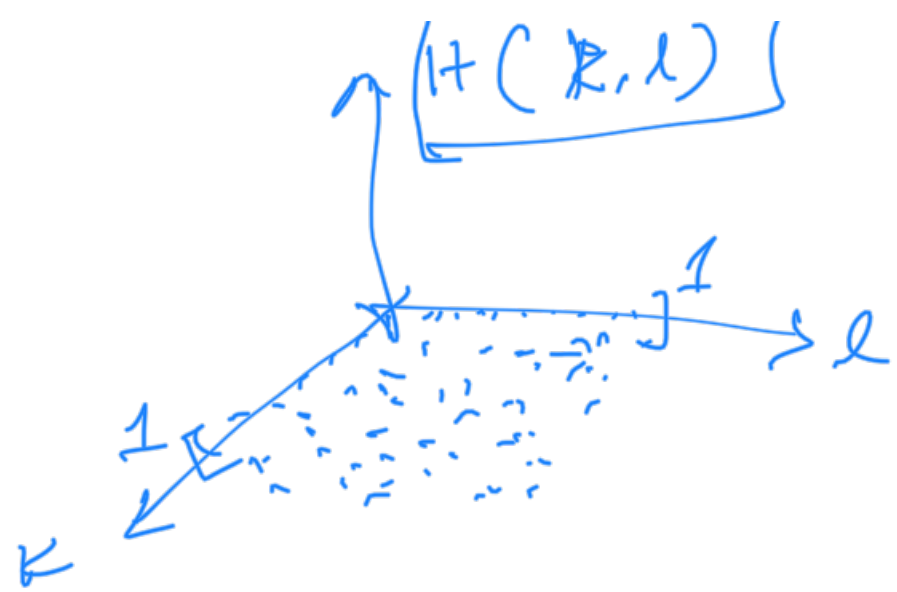
$$R_X = \{5, 2\}.$$

$$X(H) = \boxed{5}$$
$$X(T) = \boxed{2}$$

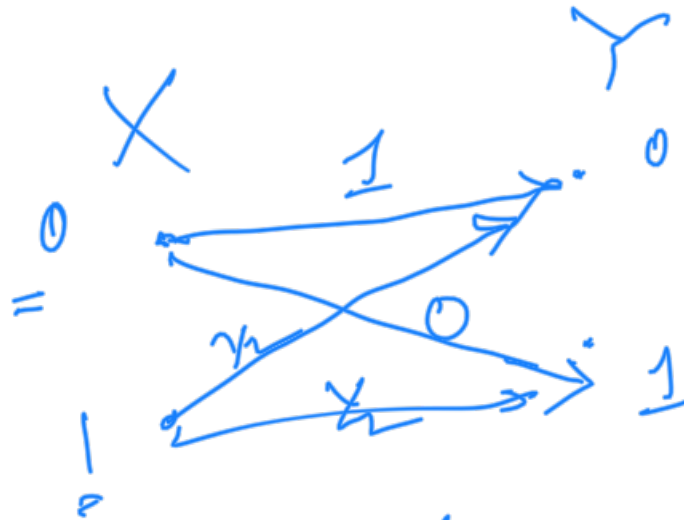
$I(X; Y)$

is related to $I(x; y)$,
 $(x, y) \in R_X \times R_Y$

Q. 10.



Q. 11

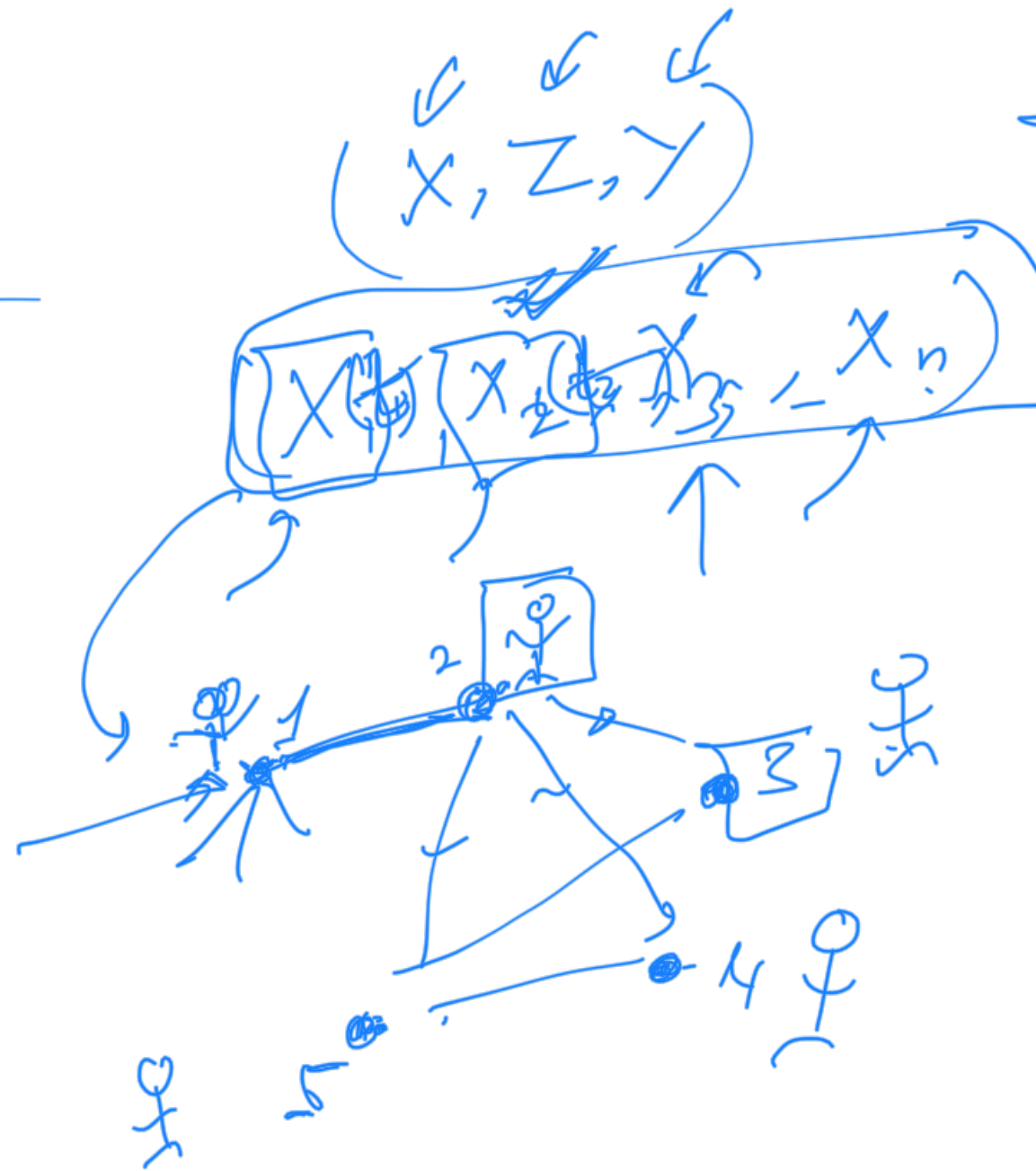


$$Q = \begin{matrix} X \rightarrow 0 & 1 \\ 1 & \end{matrix} \begin{bmatrix} \boxed{1} & \boxed{0} \\ \boxed{y_2} & \boxed{y_2} \end{bmatrix} \leftarrow \text{col. indic.}$$

$$I = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{bmatrix} p(x_1|x_1) & p(x_2|x_1) & \dots & p(x_n|x_1) \\ p(x_1|x_2) & p(x_2|x_2) & \dots & p(x_n|x_2) \\ \vdots & \vdots & \ddots & \vdots \end{bmatrix} \rightarrow 1$$

$x_m \{ p(x_1 | x_m) p(x_2 | x_m) \dots p(x_n | x_m) \}$
 row- stochastic matrix.

Q. 12.



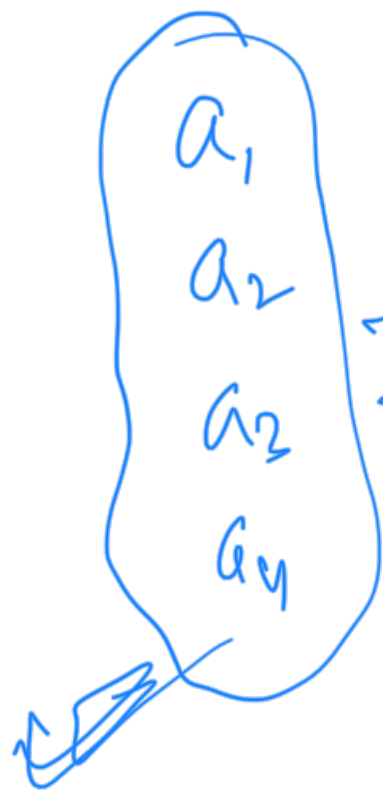
Markov chain.

$$\begin{array}{c}
 X, Y \\
 \nearrow \quad \nwarrow \\
 H \text{ --- entropy.} \\
 \boxed{I(X;Y)} \xrightarrow{\geq 0} I(X;Y) \\
 \hookrightarrow D(p(x) \parallel q(x)) \geq 0
 \end{array}$$

Operational meaning.

$$\{p(a_1), \dots, p(a_k)\}$$





$a_1 a_2 a_3 a_4 a_n$
 $a_3 a_2$
 $\vdots$

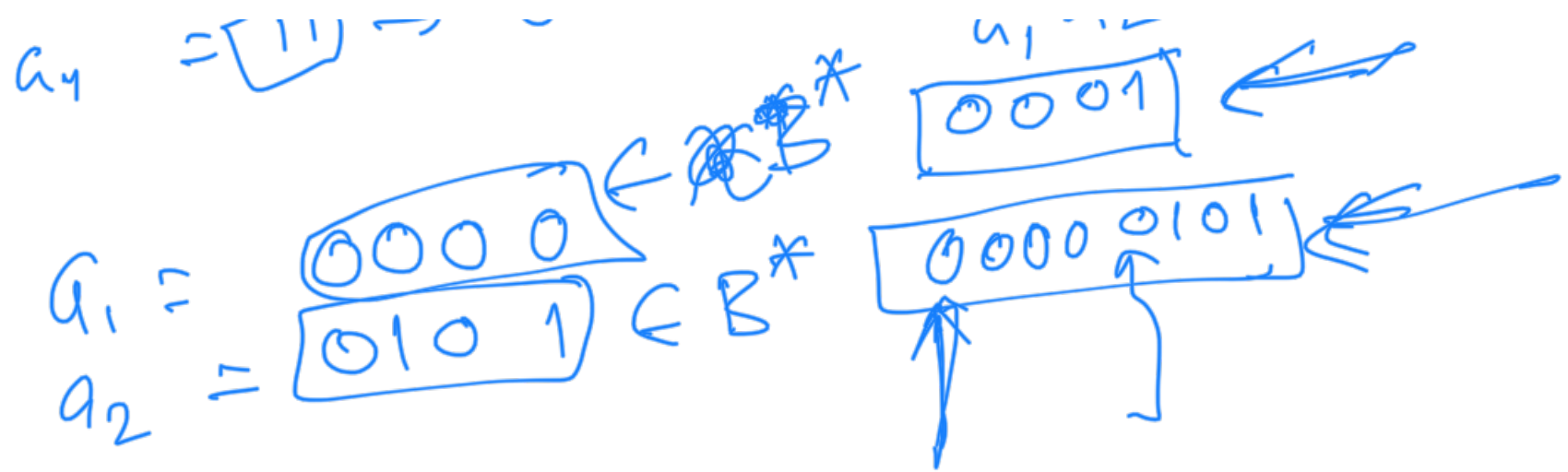
$$X = \{ a_1, \dots, a_k \} \leftarrow \text{alphabet}$$

X^* $\rightarrow$ set of strings of
 a_i 's of finite
 length.

$$x \in X$$

$a_1 = \boxed{00} \xrightarrow{C(a_1)} C(x) \rightarrow \text{codeword.}$
 $a_2 = \boxed{01} \xrightarrow{C(a_2)} C(a_2)$
 $a_3 = \boxed{10} \xrightarrow{C(a_3)} C(a_3)$
 $\vdots$
 a_n

$B = \{0, 1\}$



$$\begin{array}{ll}
 a_1 = 0 & a_2 = 0 \\
 a_2 = 1 & a_4 = 1
 \end{array}$$

Length.
 # of bit in a codeword
 $l(x) \equiv l(c(x))$

Code

$$C = \{ c(x) \mid x \in R_x \}$$

$$C = \{ c(a_k) \mid a_k \in X \}$$

Goal:

$$p(a_1)$$

$$< p(a_2) < p(a_n) < \frac{p(a_3)}{1}$$



$$1 = 1/10$$

$$1/20 = 1/30 = 1/2$$

$$\frac{1}{10} = .4$$

Expected length of a code:

Let C denote the code

know.

~~Ex~~

$$L(C) = \sum_{x \in X}$$

$$p(x) l(x)$$

minimum

$$X = \{a_1, a_2, a_3, a_4\}$$

$$\begin{aligned} p(a_1) &= \frac{1}{2} = .5 \\ p(a_2) &= \frac{1}{4} = .25 \\ p(a_3) &= \frac{1}{8} = .125 \\ p(a_4) &= \frac{1}{8} = .125 \end{aligned}$$

$$c(a_1) = [0]$$

$$c(a_2) = [10]$$

$$c(a_3) = [110]$$

$$c(a_4) = [111]$$

$$H(X) = - \sum_{x \in X} p(x) \log_2 p(x)$$

$$= - \frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{4} \log_2 \frac{1}{4} - \frac{1}{8} \log_2 \frac{1}{8} - \frac{1}{8} \log_2 \frac{1}{8}$$

$$= \left[\frac{1}{2} + \frac{1}{2} + \frac{3}{8} + \frac{3}{8} \right]$$

$$= 2.0$$

$$= \frac{14}{8} = \frac{7}{4} = \boxed{1.75 \text{ bits}}$$

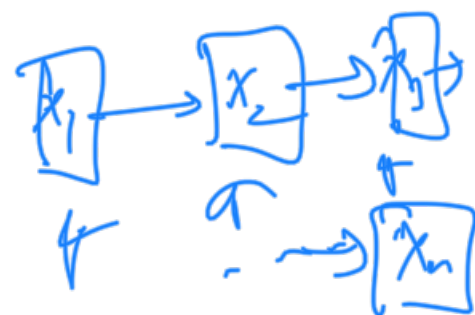
$$L(c) \leq H(X)$$

$$\begin{aligned} L(c) &= \sum p(x) l(x) \\ &= \left(\frac{1}{2} \times 1\right) + \left(\frac{1}{4} \times 2\right) \\ &\quad + \left(\frac{1}{8} \times 3\right) + \left(\frac{1}{8} \times 3\right) \\ &= \frac{1}{2} + \frac{1}{2} + \frac{3}{8} + \frac{3}{8} \\ &= 1.75 \text{ bits.} \end{aligned}$$

chain rule of entropy:

$$\bar{X} \equiv (X_1, X_2, \dots, X_n)$$

with joint $p(x_1, x_2, \dots, x_n)$ ✓



Then

$$H(X_1, X_2, \dots, X_n) = H(\bar{X})$$

$$= \sum_{i=1}^n H(X_i | X_{i-1}, \dots, X_1)$$

$$H(X, Y, Z) = \underbrace{H(X)} + \underbrace{H(Y|X)} + \underbrace{H(Z|X, Y)}$$

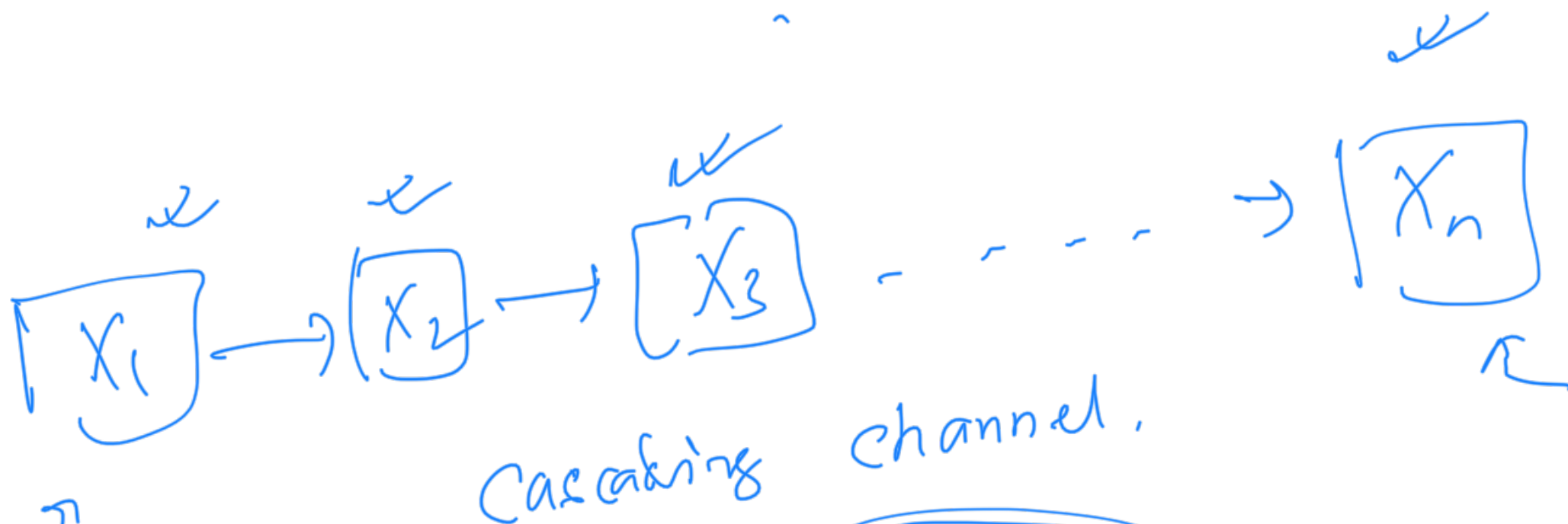
$$H(X, Y) = H(X) + H(Y|X)$$

$$H(X_1, X_2) = H(X_1) + H(X_2|X_1)$$

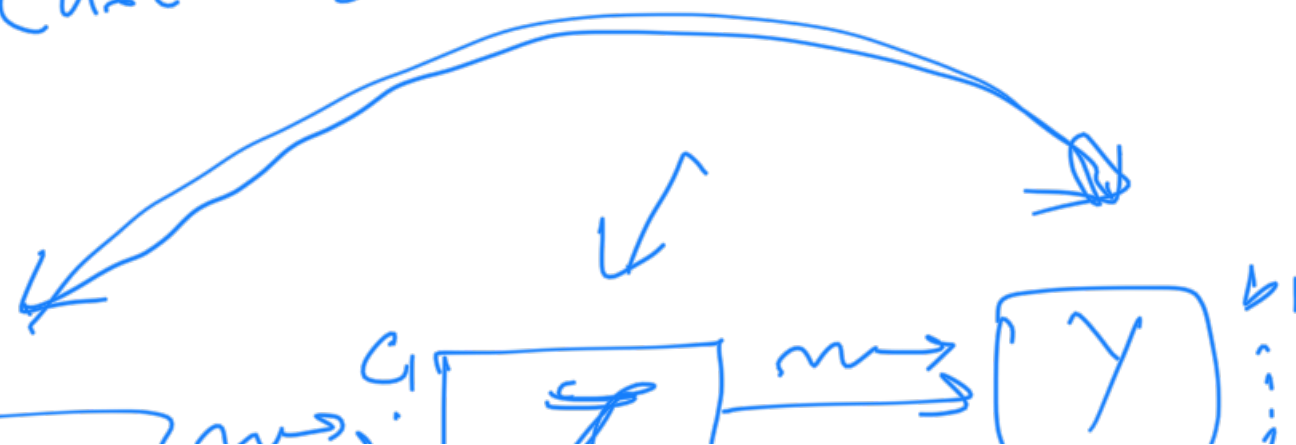
$$H(X_1, X_2) = H(X_1) + H(X_2|X_1)$$

$$H(X_1, X_2, X_3) = H(X_1) + \underbrace{H(X_2, X_3|X_1)}_{H(X_2|X_1) + H(X_3|X_2, X_1)}$$

$$\vdots$$



Proposition:





recall: $I(X; Y | Z) = H(X | Z) - H(X | Y, Z)$ ✓

$I(X; Y) = H(X) - H(X | Y)$ ✓

Chain rule mutual information.

$$I(X_1, X_2, \dots, X_n; Y) = \sum_{i=1}^n I(X_i; Y | X_1, \dots, X_{i-1})$$

Pf.

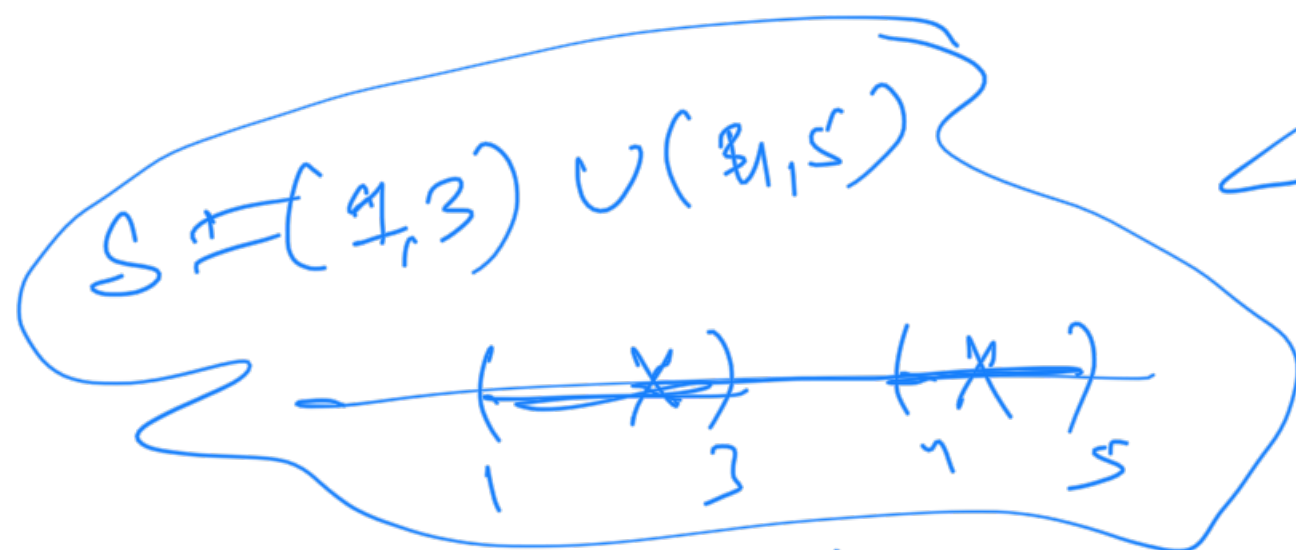
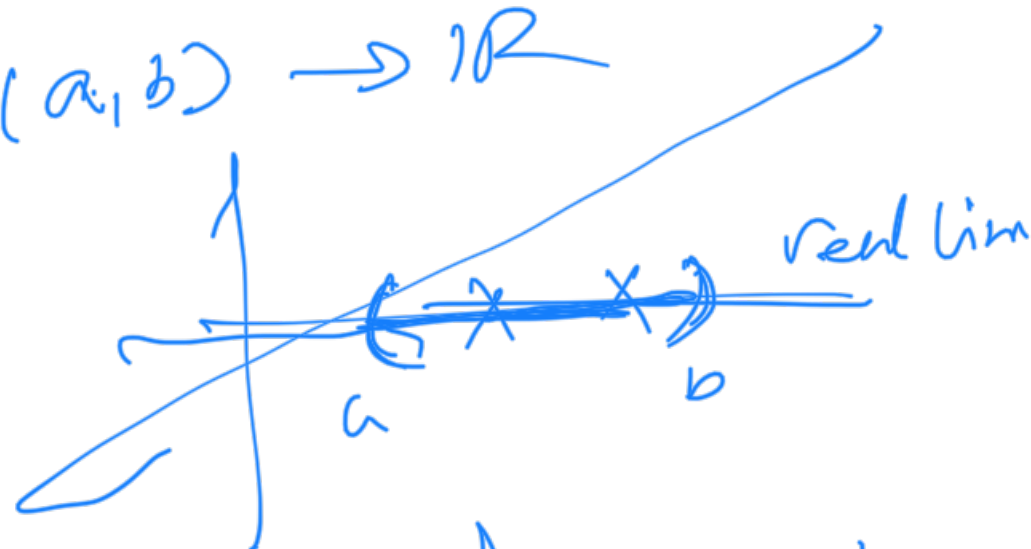
as above.

$$I(X; Y) \geq 0$$

~~$$D(p(x) || q(x)) \geq 0$$~~

Convex:

$$f: (a, b) \rightarrow \mathbb{R}$$

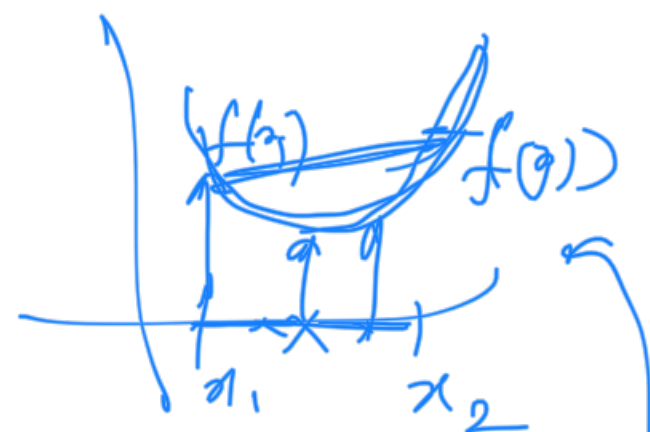


is convex if

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$$

$$0 \leq \lambda \leq 1$$

Concave: If $-f \equiv -f(x)$ is convex then f is



called concave.

[Thm. If ~~if~~ $f''(x) \geq 0$
 $\forall x \in (a, b)$ then f is convex.]

Jensen's Inequality,

$$f: () \rightarrow ()$$
$$f(x) = x^2$$

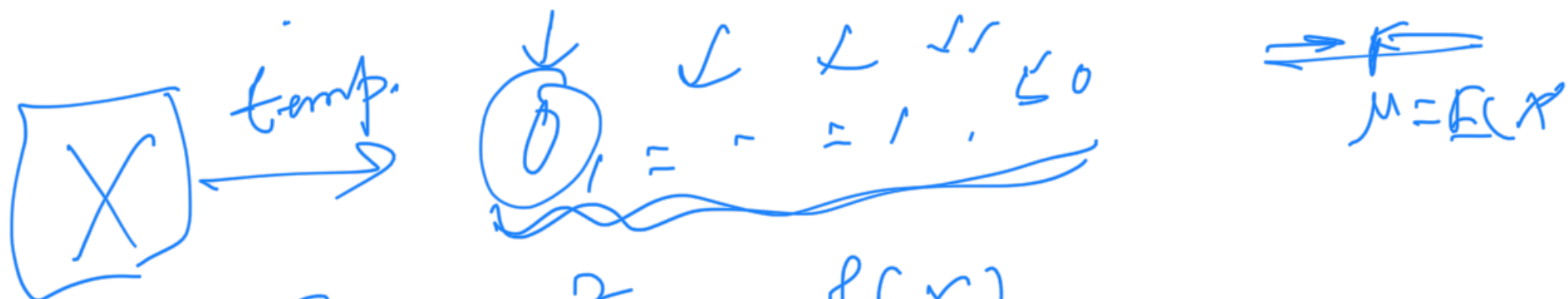
$$X \mapsto f(X) = X^2$$

If f is a convex function and
 X is a random variable then

$$E[f(X)] \geq f[E(X)] !!!$$



$$E(Y) = \sum_{y \in \mathcal{Y}} y p(y)$$



$$\mu = E(X)$$

$$E(X) = Y = X - E(X)^2 = f(X)$$

Prop.

$$D(p \parallel q) \geq 0$$

$$-D(p \parallel q) \leq 0$$

Pf.

$$-D(p \parallel q) = - \sum p(x) \log \frac{p(x)}{q(x)}$$

H.W.
 $\log$ is concave!!

$$\rightarrow \boxed{\leq} \boxed{\log} \underbrace{\sum_x p(x) \frac{q(x)}{p(x)}}_{\leftarrow} \leftarrow$$

$$= \log [1]$$

$$= 0$$

H.w.

$$D(p||q) = 0$$

iff $p = q \leftarrow$

Corollary:

$$\underbrace{I(X; Y)}_{\parallel} \geq 0.$$

$$D(p(x, y) \parallel p(x) p(y))$$

Upper bound for the entropy:

$$H(X) \leq ??$$

Suppose:

$X \rightarrow$ random variable on

the sample space $S = \{ \text{~~all~~ } \{0,1\}^m \}$
 $= \{ \underbrace{x_1 x_2 \dots x_m}_{x_i \in \{0,1\}} \}$

$|\{0,1\}^m| = 2^m$

$p(x) = \frac{1}{2^m} = \boxed{2^{-m}}$

$H(X) = - \sum_{x \in S} p(x) \log p(x)$
 $= \boxed{m \text{ bits}}$

$\boxed{100}$

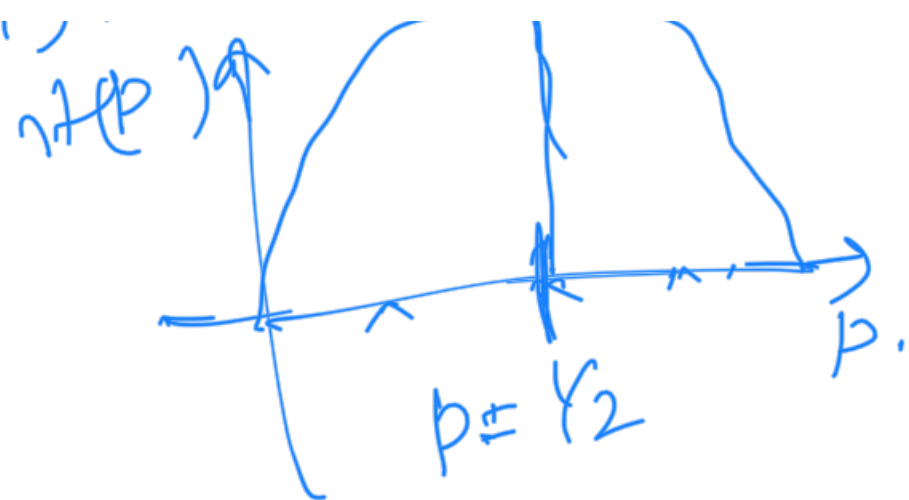
$C = \{ H, T \}$
 $p(H) = 1-p$

~~max!!~~

?

$$p(H) = p, \quad p(T) = 1-p$$

$$p(H) = \frac{1}{2} = p(T)$$



$$H(X) \leq \boxed{?}$$

$$A^* = \{0, 1\}$$

Lec-6

$\hat{a}_1$
 $\hat{a}_2$

a_{2^n}



$$p(a_k) - 1$$

?

$$p(a_1), \dots, p(a_n)$$

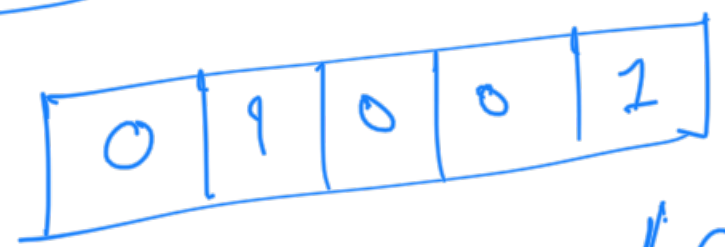
$$S = \{0, 1\}^n = \{ \underline{x_1, x_2, \dots, x_n} \mid x_i \in \{0, 1\} \}$$

pmf.

$$p(x) = \frac{1}{2^n} = 2^{-n}, \quad x \in S$$

$$H(x) = n \text{ bits.}$$

??



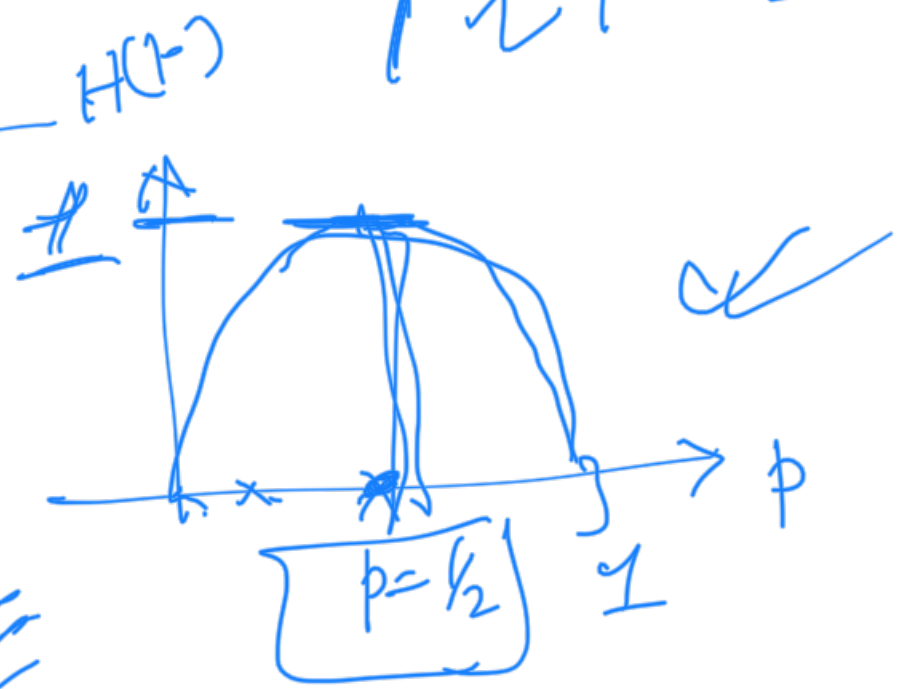
2^5

$$n = 5$$

$$|S| = 2^n$$

$$S = \{H, T\}$$

$$\begin{aligned} p(H) &= p \\ p(T) &= 1-p \\ p &= 1/2 \end{aligned}$$



$$H(x) \leq ??$$

5

Prop.

$$H(X) \leq \log |X|$$

l.u.b.

Pf. Suppose $u(x)$ denotes the uniform pmf on X .
 Let $p(x)$ be any pmf on X

$$0 \leq D(p \parallel u) = \sum_{x \in X} p(x) \log \frac{p(x)}{u(x)}$$

$$= \sum_{x \in X} p(x) \log p(x) - \sum_{x \in X} p(x) \log u(x)$$

$$\stackrel{1}{=} -H(X) \rightarrow \sum_{x \in X} p(x) \log \frac{1}{p(x)}$$

$$\stackrel{2}{=} \log |X| - H(X)$$

$$\Rightarrow -H(X) \geq -\log |X| \quad \checkmark$$

$$\Rightarrow H(X) \leq \log |X|.$$

Prop.

$$H(X|Y) \leq H(X) \quad !!!$$

the equality occurs when X & Y are independent.

are stochastic

$$\text{Pf.} \quad 0 \leq I(X; Y) = \underbrace{H(X)} - H(X|Y)$$

$$\Rightarrow H(X) \geq \underline{H(X|Y)}.$$

Prop. $\bar{X} = (X_1, X_2, \dots, X_n) \leftarrow$

proof $(X_1, X_2, \dots, X_n)$

$a_1,$
 $a_2,$
 $\vdots$
 a_k

$\rightsquigarrow$

$$(a_i, a_j, a_k) \equiv (X_1, X_2, X_3)$$

$$\underline{a_i, a_j, a_k} \quad X_i \sim X$$

Q

$$H(X) \leq \sum_{i=1}^n H(X_i)$$

$$H(X_1, X_2, \dots, X_n) = \sum_{i=1}^n H(X_i | X_1, \dots, X_{i-1})$$

and it is equality if X_i 's are iid

Pr. $H(X_1, \dots, X_n) \leq \sum_{i=1}^n H(X_i | X_1, \dots, X_{i-1})$

$$\leq \sum_{i=1}^n H(X_i)$$

$p(a_1), p(a_2), \dots, p(a_K)$

proof.

$\{a_1, a_2, \dots, a_K\}$

$\phi_{ik} = p(a_k)$

$\phi_i \geq 0, \sum_{i=1}^K \phi_i = 1$

$S = \left\{ \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_K \end{bmatrix} \in \mathbb{R}^K \right\}$

Set of
all p.m.f.s
on $\{a_1, \dots, a_K\}$.

$\subseteq$

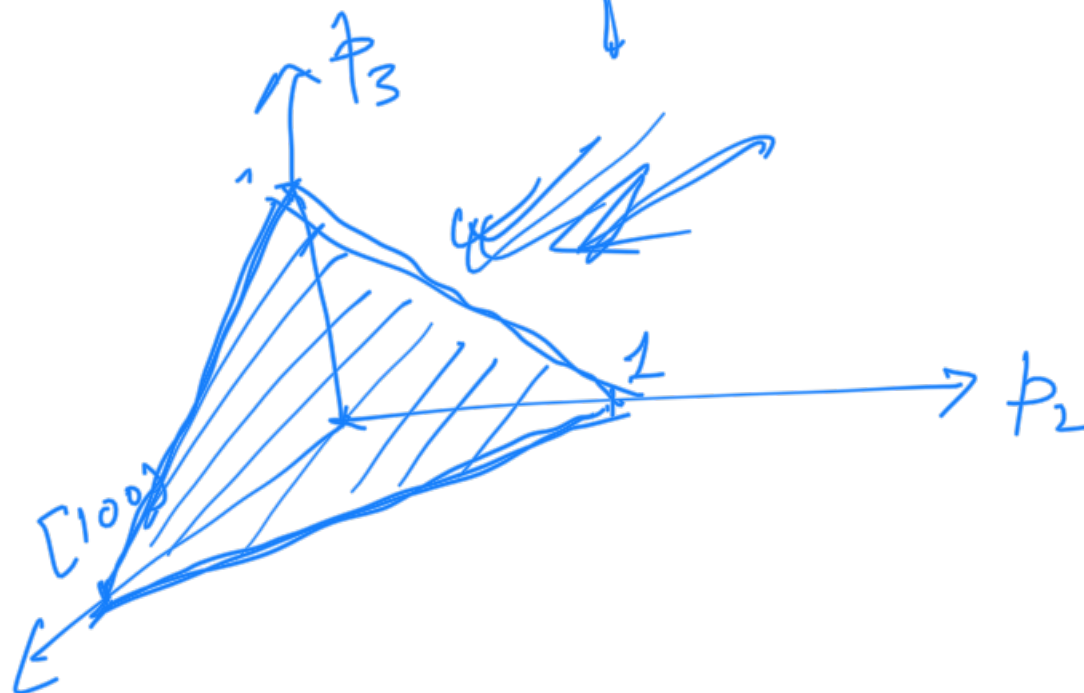
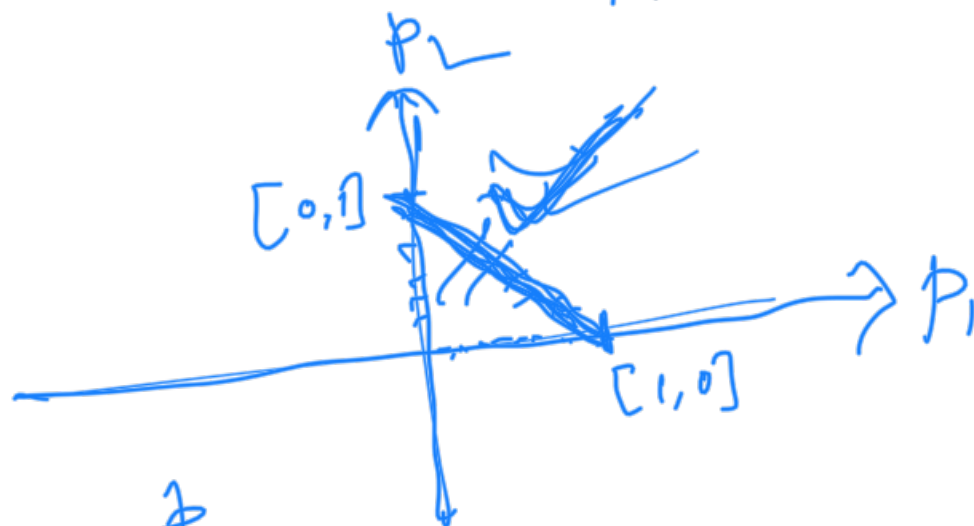
$K=2$

$K=3$

H.W.

$\mathbb{R}^K$

$$\mathbb{R}^K = \underbrace{\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}}_{K\text{-times.}}$$



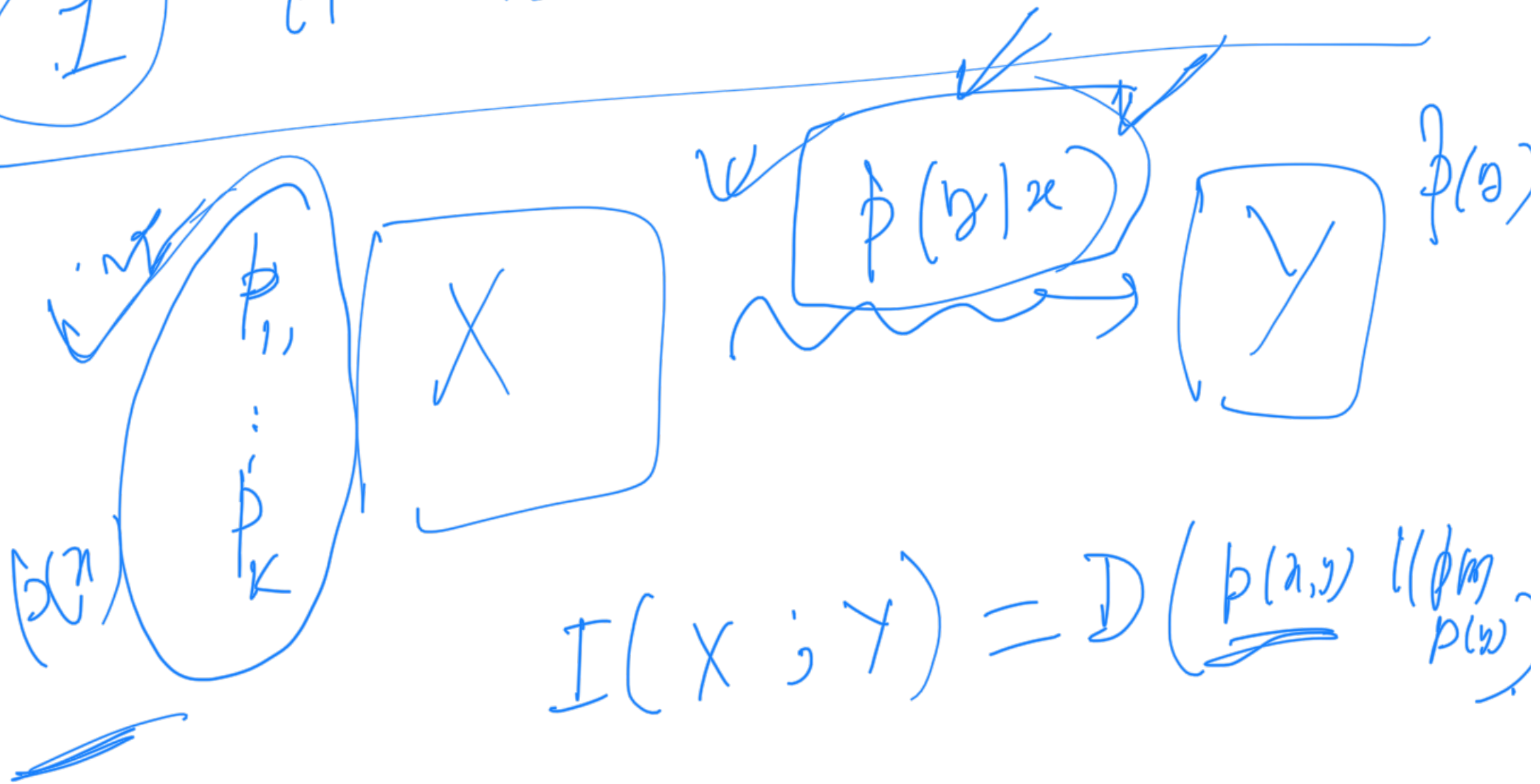
p_1

$$\Delta v = b$$

$$A = \{a_1, \dots, a_K\} \quad (1 \wedge \dots)$$

$$H: \mathbb{S} \rightarrow \mathbb{R}[0, \alpha)$$

① H is a concave function.



$$I(X; Y) = D(\underline{p(x, y)} \| p(x) p(y))$$

1 $I(X; Y)$ is concave function
w.r.t $p(x)$ is fixed.

2 $I(X; Y)$ is convex function
w.r.t $p(y|x)$ is fixed.



pmix

