ETE 311 26-11-24

(*) Thomas A cover, Ranjan Bose.

(*) Probability → 100% → message

less than 100% → Information.

data compression ← Source coding, channel coding

Discrete source → Finite num of ele. produce
-ছবি (Audio/Video ...)

Memoryless → Present value, past value -এর
-উপর dependable না.

D. Memory. source → dependent.

(*) ~~Higher~~ Lower probability implies a higher
degree of uncer, where ~~more~~, a higher ~~prob.~~ uncert.
contains more/higher info.

Consider a source of fixed ele. $x, x_i \in X$
where $i = 1, 2, \ldots n$ with probability
distribution P_i & $\sum_{i=1}^{n} P_i = 1$. self info $I(x_i) =$

$$\log_2\left(\frac{1}{P_i}\right)$$

$$= -\log_2 P_i$$

Let, $x_1 = \{x_1, x_2, x_3\}$.

$P = \{0.5, 0.25, 0.25\}$.

$\boxed{P \propto \frac{1}{I(x)}}$

$\therefore I(x_1) = \log_2\left(\frac{1}{0.5}\right) = 1$ bit.

$I(x_2) = \log_2\left(\frac{1}{0.25}\right) = 2$ bit.

~~$I(x_3) =$~~

Lower probability requires higher space in terms of bit

Higher — Lower

Self info & properties:

⇒ i) If we are absolutely certain of the outcome of an event, even before it occure there is no info gained. $I(x_i) = 0$; if $P_i = 1$ means certainity message

ii) The occurance of an event either provides some or no info, but never brings about a loss of info. $I(x_i) \geq 0$, for $0 \leq P(x_i) \leq 1$.

iii) The less probable an event is, the more info we gain when it occures $I(x_1) > I(x_2)$ if $P_1 < P_2$

iv) $I(x_1, x_2) = I(x_1) + I(x_2)$ x_1 & x_2 → statistically indepe.

$= \log\left(\frac{1}{P_1}\right) + \log\left(\frac{1}{P_2}\right)$

$= \log\left(\frac{1}{P_1 P_2}\right) = \log\left(\frac{1}{P(1,2)}\right)$

(*) Entropy → avg. num. of bit per sym. represent করে,

এই text → a b c.

Let, $P(a) > P(b) > P(c)$.

$\downarrow$ 2 bit (min) $\downarrow$ 4 bit $\downarrow$ 8 bit

→ এর নিচে কোনো message ~~fix~~ হবে না (boundary)

⇒ Measurement of uncertainity of any event can be defined as $H(x_i) = -P_i \log_2 P_i$ where $x_i \in X$ & $\sum P_i = 1$

$$H(x) = \sum_{x_i \in X} H(x_i) = -\sum_{i=1}^{n} P_i \log_2 P_i \text{ bits/sym.}$$

Ex $P_1 = 0.1$, $P_2 = 0.2$, $P_3 = 0.3$, $P_4 = 0.4$

$$\Rightarrow H = -\sum_{i=1}^{4} P_i \log_2 1/P_i = -\left(0.1 \log_2 1/0.1 + 0.2 \log_2 1/0.2 + 0.3 \log_2 1/0.3 + 0.4 \log_2 1/0.4\right)$$

$$= 1.8464 \text{ bits/sym (min bit)}$$

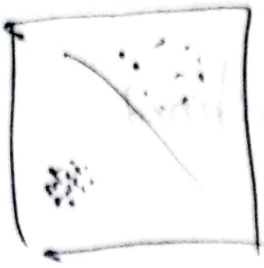
$$\approx 2 \text{ bits/sym.}$$

Self Info

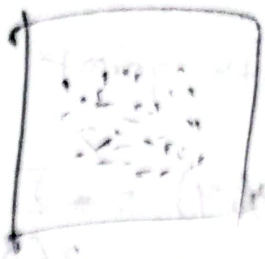
(*) Unit not only bit but also Hartley, nat.

Ass.

for purchase. 1-12-24



fraud
real/genuine.



Low entropy

High entropy
equal probability.

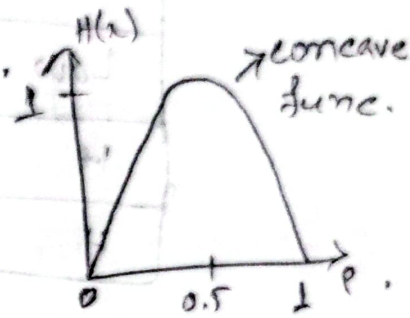
BPSK : $x = \begin{cases} 1, \ldots P \\ 0, \ldots (1-P) \end{cases}$

range 1 এর মধ্যে সম্ভাবনা

$$H(x) = -\sum_{x=0}^{1} P(x) \log_2 P(x) = -P \log_2 P - (1-P) \log_2 (1-P)$$

when $P = 0, 1$ there is no entropy.

$P = 0.5$, $H(x) = 1$ (max)



(*) $P(00) = \frac{2}{15}$ $P(01) = \frac{4}{15}$

$P(10) = \frac{6}{15}$ $P(11) = \frac{3}{15}$.

$$H(X) = -P(00) \log_2 P(00) - P(10) \log_2 P(10) - P(01) \log_2 P(01) - P(11) \log_2 P(11)$$

$$= -\frac{2}{15} \log_2 \frac{2}{15} - \frac{6}{15} \log_2 \frac{6}{15} - \frac{4}{15} \log_2 \frac{4}{15} - \frac{3}{15} \log_2 \frac{3}{15}$$

$= 1.889$ bits/symbol.

Joint entropy: $H(x,y)$ of a pair of random variable

(x,y) with joint distribution $P(X,Y)$ is defined as

$$H(X,Y) = -\sum_{x\in X}\sum_{y\in Y} P(x,y)\log_2 P(x,y)$$

$$= -E\log_2 P(x,y)$$

Studied Hour (X)	Attendance (Y)	$P(x,y)$	
Low	Low	0.2	$\rightarrow H_1 = -0.2\log_2 0.2 = 0.46$
Low	High	0.1	$\rightarrow H_2 = -0.1\log_2 0.1 = 0.33$
H	L	0.15	$\rightarrow H_3 = -0.15\log_2 0.15 = 0.41$
H	H	0.55	$\rightarrow H_4 = -0.55\log_2 0.55 = 0.47$

$$\sum_{n=1}^{4} H_n = 1.67.$$

$$H(x) = -P(X=\text{Low})\log_2 P(X=\text{Low}) - P(X=\text{high})\log_2 P(X=\text{high})$$

Low $= 0.2 + 0.1 = 0.3$

High $= 0.15 + 0.55 = 0.7$

$$= -0.3\log_2 0.3 - 0.7\log_2 0.7.$$

$$= 0.883$$

$H(x,y)$ dependent.

$$H(y) = 0.934$$

Mutual info.

$$H(x) + H(y) = 1.817.$$

2টি var. obs. করলে অপর var. এর সম্বন্ধে জানা যায়

Mutual Info: $I(x,y) = H(x) + H(y) - H(x,y)$.

$Y \backslash X$	0	1
0	1/2	1/4
1	1/8	1/8

3-12-24

⊛ Conditional entropy:

$H(Y|x)$ → আগেরটা X জেনে ছিলাম then Y.

→ Y এর uncertainity কেমন হবে after observing x.

⊛ dependent হতে হবে।

Conditional entropy of Y given X as follows:

$$H(Y|x) = \cancel{-\sum P(x) \sum P(y|x)}$$ → basically prob. of y.

$$= -\sum_x P(x) \sum_y P(y|x) \log_2 P(y|x)$$

$$= -\sum_{x \in X} \sum_{y \in Y} P(x,y) \log_2 P(y|x)$$

$$P(x|y) = \frac{P(x,y)}{P(y)}$$

$$H(x,y) = -\sum_{x \in X} \sum_{y \in Y} P(x,y) \log P(x,y)$$

$$= -\sum_{x \in X} \sum_{y \in Y} P(x,y) \log P(x) P(y|x)$$

$$= -\sum_{x \in X} \sum_{y \in Y} P(x,y) \log P(x) - \sum_{x \in X} \sum_{y \in Y} P(x,y) \log P(y|x)$$

$$= -\sum_{x \in X} P(x) \log P(x) - \sum_{x \in X} \sum_{y \in Y} P(x,y) \log P(y|x)$$

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

* $H(x|y) = -\sum_{xy} \sum P(xy) \sum_{x} P(x|y) \log_2 (x|y)$.

$= -\sum_{y \in Y} \sum_{x \in X} P(x,y) \log_2 (x|y)$.

* $H(x,y) = -\sum_{y \in Y} \sum_{x \in X} P(x,y) \log P(x,y)$.

$= -\sum_{y \in Y} \sum_{x \in X} P(x,y) \log P(y) \cdot \log P(x|y)$

$= -\sum_{y \in Y} \sum_{x \in X} P(x,y) \log P(y) + \sum_{y \in Y} \sum_{x \in X} P(x,y) \log P(x|y)$

$= -\sum_{y} P(y) \log P(y) - \sum_{y \in Y} \sum_{x \in X} P(x,y) \log P(x|y)$

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

actual value, predicted → Relative ent. → KL Divergence.

* $D(p||q) = \sum P(x) \log_2 \frac{p(x)}{q(x)}$ → True value, → estimated value.

* if $p = q$, $D(p||q) = 0$.

~~inefficiency of the~~

Relative ent. $\propto \frac{1}{\text{prediction}}$

~~bit~~

Microprocessor

8-12-24

$$I(x;y) = \sum_{x,y} P(x,y) \log \frac{P(x,y)}{P(x)P(y)} \rightarrow D(P(x,y) \| P(x)P(y))$$

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

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

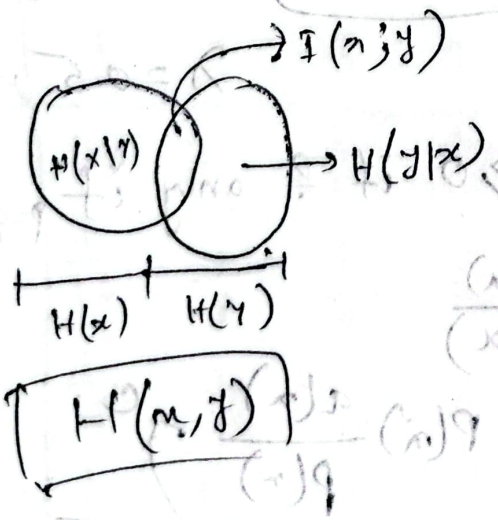
$$= -\sum_{x,y} P(x,y) \log P(y) + \sum_{x,y} P(x,y) \log P(y|x)$$

$$= H(y) - H(y|x) \quad \text{—— (i)}$$

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

$$H(y|x) = H(x,y) - H(x) \quad \text{—— (ii)}$$

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



If the function f has a 2nd derivative that is non-negative over an interval, the function is convex.

$$\frac{d^2 f(x)}{dx^2} \geq 0 \quad f(x) \text{ is convex} \quad \text{Ex: } x^2, e^x$$

$f(x) = x^2 \quad f(0) = 0 \quad f(2) = 4$ avg. of the func $= \frac{0+4}{2} = 2$

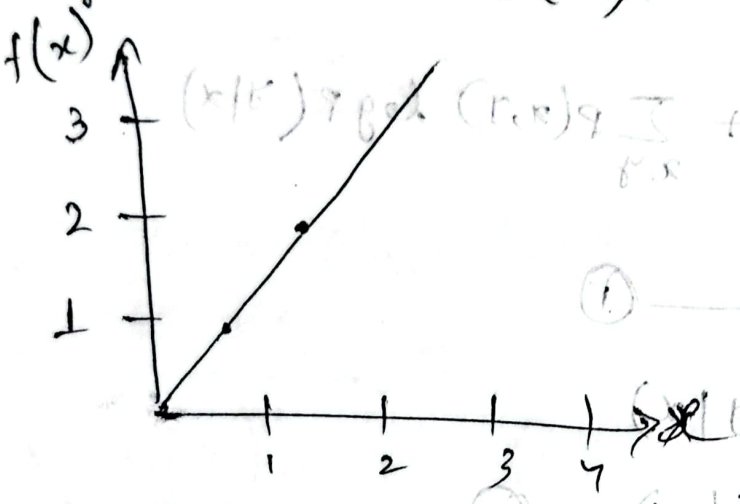
func. of the avg:

avg of the func $\geq f(\text{avg})$

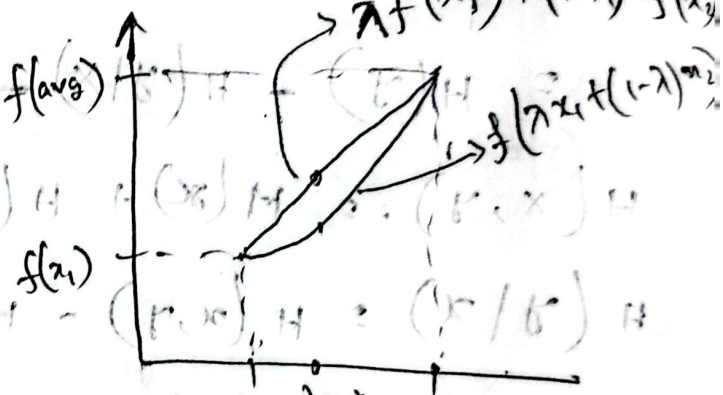
$a = 0, \; b = 2$

$\text{avg} = \frac{0+2}{2} = 1.$

$f(\text{avg}) = 1^2 = 1$



$E f(x) < f(E(x))$



$$\boxed{\lambda f(x_1) + (1-\lambda) f(x_2) \geq f(\lambda x_1 + (1-\lambda) x_2)}$$

$2 \geq 1$

$x_1 = 0 \quad x_2 = 2$

$\lambda = 0.5.$

<u>Gibbs info Inequility</u>: $D(p \| q) \geq 0$ if & only if $p = q$

$$-D(p \| q) = -\sum_{x \in X} p(x) \log \frac{p(x)}{q(x)}$$

$$= \sum_{x \in X} p(x) \log \frac{q(x)}{p(x)} \leq \log \sum_{x \in X} p(x) \frac{q(x)}{p(x)}$$

$$= \log \sum_{x \in X} q(x) \leq \log 1 = 0 \quad \left[E\{f(x)\} \geq f(E\{x\}) \right]$$

$D(p\|q) \neq D(q\|p)$

(8) $X = \{0, 1\}$ $\overbrace{p(0) = 1-r,\ p(1) = r}^{\text{actu.}}$, ~~$s(0) =$~~

$\underbrace{q(0) = 1-s \quad q(1) = s}_{\text{rec.}}$. $r = 0.5 \quad s = 0.25$.

$$D(p\|q) = \sum p(x) \log_2 \frac{p(x)}{q(x)}$$

$$= p(0) \log_2 \frac{p(0)}{q(0)} + p(1) \log_2 \frac{p(1)}{q(1)}$$

$$= 0.5 \log_2 \frac{0.5}{0.75} + 0.5 \log_2 \frac{0.5}{0.25} = 0.$$

$$= 0.2075.$$

$$= (1-r) \log_2 \frac{1-r}{1-s} + r \cdot \log_2 \frac{r}{s} = 0.$$

$$D(q\|p) = 0.75 \log_2 \frac{0.75}{0.5} + 0.25 \log_2 \frac{0.25}{0.5}$$

$= 0.1887$. [Asymetry]

Common

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

$y \backslash x$	1	2	3	4
1	$1/8$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$
2	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$
3	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
4	$\frac{1}{4}$	0	0	0

$$H(x) = \sum_{x \in X} P(x, y) \log_2 \frac{1}{P(x)}$$

$$= \sum_{x=1}^{4} = P(x_1) \log_2 \frac{1}{P(x_1)} + P(x_2) \log_2 \frac{1}{P(x_2)} + P(x_3) \log_2 \frac{1}{P(x_3)} + P(x_4) \log_2 \frac{1}{P(x_4)}$$

$$P(x_1) = \frac{1}{8} + \frac{1}{16} + \frac{1}{16} + \frac{1}{4} = \frac{1}{2}$$