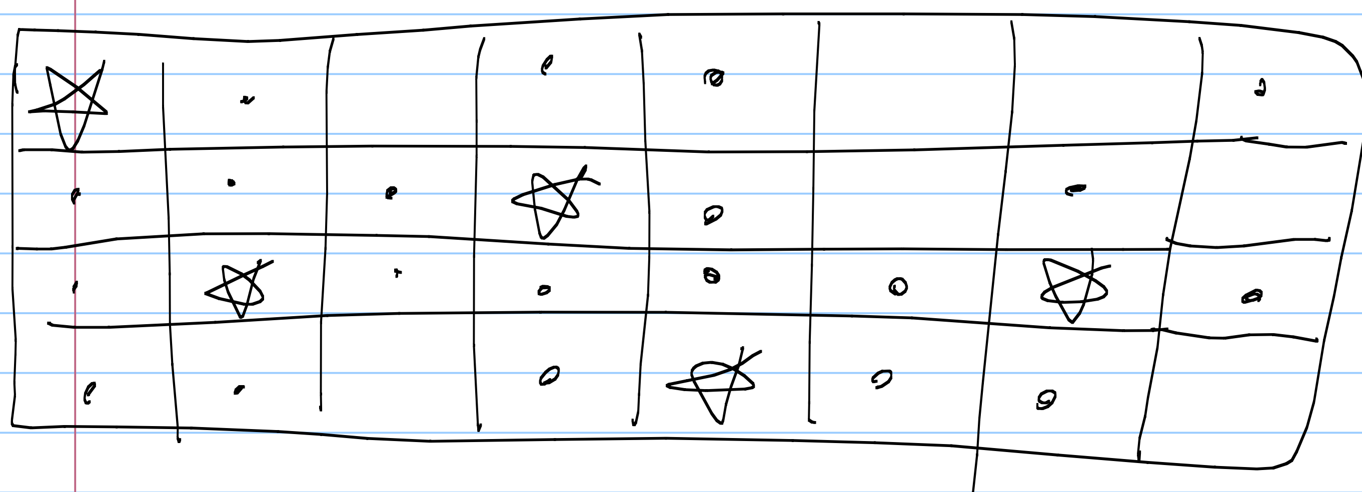
Lecture 05

Note Title 1/25/2012

* The connection between divergence & entropy

* The connection between divergence & channel capacity

  - Sphere packing bound.



$$\log_2 5 = 2.3219$$

upper bound $\log \frac{32}{5} = 2.678$

$$\text{rate}: \frac{\log_2(5)}{\log_2(32)} = \frac{2.3219}{5} \approx 0.46$$

Example 2:

a bit string of n bits 010110...

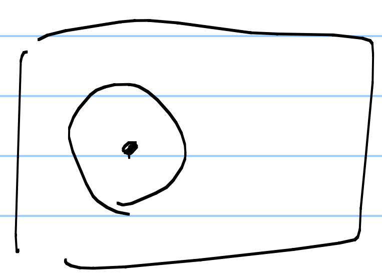
If noise-free, 2^n possibilities

$$= \log_2(2^n) = n \text{ bits.}$$

A special noise model: At most $0.1n$ bits are flipped (corrupted)

Q: What is the achievable rate upper bound?

A: each choice occupies



$\sum_{k=0}^{0.1n} \binom{n}{k}$ different bit strings

$$\Rightarrow \text{the reduce info} = \frac{2^n}{\sum_{k=0}^{0.1n} \binom{n}{k}}$$

$$= \log_2 \left(\frac{2^n}{\sum \binom{n}{k}} \right) \text{ bits}$$

In terms of bits: $\log_2\left(\dfrac{2^n}{\sum_{k=0}^{0.1n}\binom{n}{k}}\right)$ bits.

$$= n - \log_2\left(\sum_{k=0}^{0.1n}\binom{n}{k}\right)$$

The normalized rate (upper bound) becomes

$$\frac{n - \log_2\left(\sum_{k=0}^{0.1n}\binom{n}{k}\right)}{n} \triangleq r(n)$$

$r(n)$ is an increasing function

Q: $\lim_{n\to\infty} r(n) = ?$

Or equivalently $\lim_{n\to\infty} \dfrac{\log_2\left(\sum_{k=0}^{0.1n}\binom{n}{k}\right)}{n} = ?$

Ans:

$$\frac{\log_2\left(\binom{n}{0.1n}\right)}{n} \le \frac{\log_2\left(\sum_{k=0}^{0.1n}\binom{n}{k}\right)}{n} \le \frac{\log_2\left(0.1n\cdot\binom{n}{0.1n}\right)}{n}$$

By Stirling's formula $n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$

$$\frac{n!}{(0.1n)!(0.9n)!} \approx \frac{\sqrt{2\pi n}\, n^n / e^n}{\frac{\sqrt{2\pi\, 0.1n}\,(0.1n)^{0.1n}\sqrt{2\pi\, 0.9n}\,(0.9n)^{0.9n}}{e^n}}$$

$$\approx \frac{1}{\sqrt{2\pi(0.1)(0.9)n}} \frac{n^n}{(0.1n)^{0.1n}(0.9n)^{0.9n}}$$

$$\frac{\log_2 \binom{n}{0.1n}}{n} \approx \frac{n}{n}\log_2 n - \frac{(0.1n)}{n}\log_2(0.1n)$$
$$-\left(\frac{0.9n}{n}\right)\log_2(0.9n)$$

$$= -0.1\log_2(0.1) - 0.9\log_2 0.9$$

$$= H(0.1)$$

$$\frac{\log\left(0.1n \cdot \binom{n}{0.1n}\right)}{n} \approx \frac{\log(0.1n)}{n} + H(0.1)$$

$$\Rightarrow \lim_{n\to\infty} r(n) = \underline{1 - H(0.1)}$$

The Shannon capacity of a BSC with crossover prob 0.1.

The 2nd Application of $D(P_0 \| P_1)$

* Consider i.i.d channel

$$\xrightarrow{X_i} \boxed{P_{Y|X}} \xrightarrow{Y_i}$$

If the codeword satisfies that the marginal distribution of X_i is P_X then the achievable rate is

$$\leq \underbrace{I(X;Y)}_{\text{mutual information}} \triangleq \underline{E_{XY}}\left(\log \frac{P_{XY}(X,Y)}{P_X(X)\cdot P_Y(Y)}\right)$$

$\hookrightarrow$ evaluated by P_{XY}

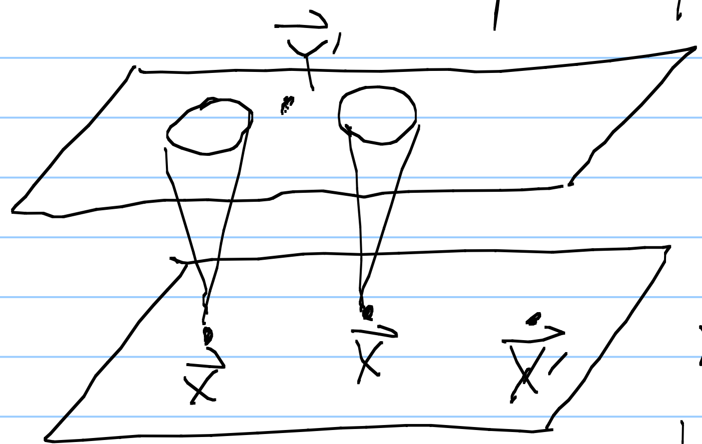
$$= D(\underline{P_{XY}} \,\|\, \underline{P_X \times P_Y})$$

look as if dependent   actual indep

H_0: $(X_1, Y_1), \cdots, (X_n, Y_n)$ are i.i.d with joint distri P_{XY}

H_1: $(X_1, Y_1) \cdots (X_n, Y_n)$ are i.i.d with independent distri $P_X \times P_Y$.

* Consider the sphere packing bound.



$Y_1, \cdots, Y_N$

$x_1, \dots, x_n$ decided by P_X

The larger the base $(x_1, \cdots, x_n)$ the more "cones" we can squeeze into.

Therefore, we need to fix $\boxed{P_X.}$

& discuss the # of cones one can fit.

→ The stronger the signal power.

Question: Consider $\vec{X}' \sim (P_X)^n$

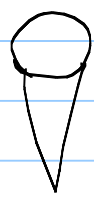
$\vec{Y}' \sim (P_Y)^n$

& $\vec{X}', \vec{Y}'$ are indep.

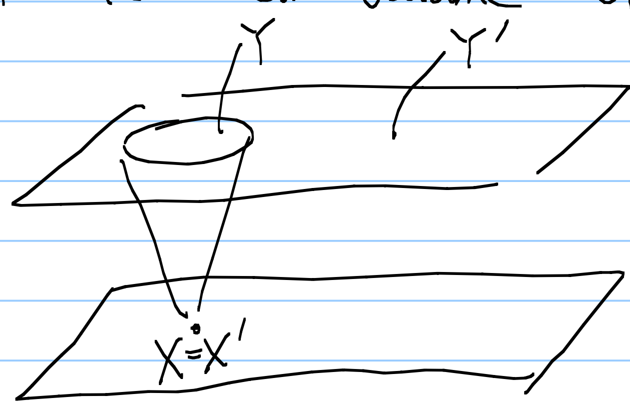
when will $\vec{X}', \vec{Y}'$ look like it is generated from $P_X, P_{Y|X}$? (or from P_{XY})

Ans: When $\vec{X}', \vec{Y}'$ fall into the

cone

$\Rightarrow$  is taking up $e^{-nD(P_0 \| P_1)}$

of the total (prob) volume of



where P_0 : P_{XY} X, Y depend on each other

P_1 : $P_X \times P_Y$ are indep

As a result, we should be able to pack

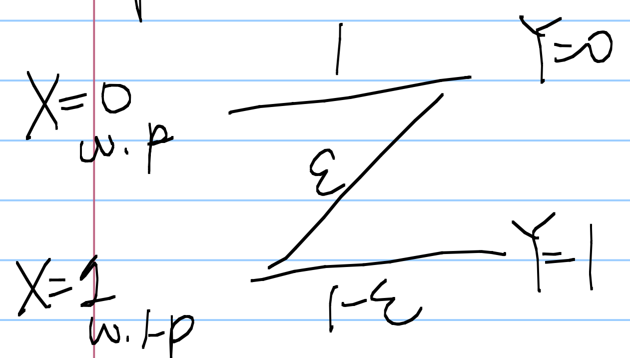
$$e^{n D(P_{XY} \| P_X \cdot P_Y)} = e^{n I(X;Y)}$$

"codewords" without significant overlap.

$$I(X;Y) = D(P_{XY} \| P_X \times P_Y)$$

$$= E_X \left(D\left(P_{Y|X}(\cdot|X) \| P_Y(\cdot) \right) \right)$$

Example: The capacity of a Z channel



Step 1: Assume the prior distribution P_X

Step 2: Evaluate $I(X;Y)$

$$= E_{XY}\left(\log \frac{P_{XY}(X,Y)}{P_X(X) \cdot P_Y(Y)} \right)$$

$(x,y) = (0,0)$ $\quad\quad\quad\quad\quad\quad$ $(1,0)$

$$I(X;Y) = p \log \frac{p}{p\cdot(p+\varepsilon(1-p))} + \varepsilon(1-p)\log\frac{(1-p)\varepsilon}{(1-p)(p+\varepsilon(1-p))}$$

$(1,1)$

$$+ (1-p)(1-\varepsilon) \log \frac{(1-p)(1-\varepsilon)}{(1-p)((1-p)(1-\varepsilon))}$$

Step 3: $\max_{p} I(X;Y)$

$$p^* = \frac{1-\varepsilon^{\frac{1}{1-\varepsilon}}}{1+(1-\varepsilon)\varepsilon^{\frac{\varepsilon}{1-\varepsilon}}}$$

Example: Consider a BPSK system over an i.i.d. channel.

I.e. $X_i = (-1)^{b_i}\cdot a$ $\quad$ b_i is the i-th bit.

$Y_i = X_i + N_i$

N_i: I.i.d. Gsn $\mathcal{N}(0,1)$

Q: Find out the optimal achivable rate of this channel.

Ans: Step 1: Design the marignal prob P_X

Say $P(X=a)=p \qquad P(X=-a)=(1-p)$

Step 2: $I(X;Y)$

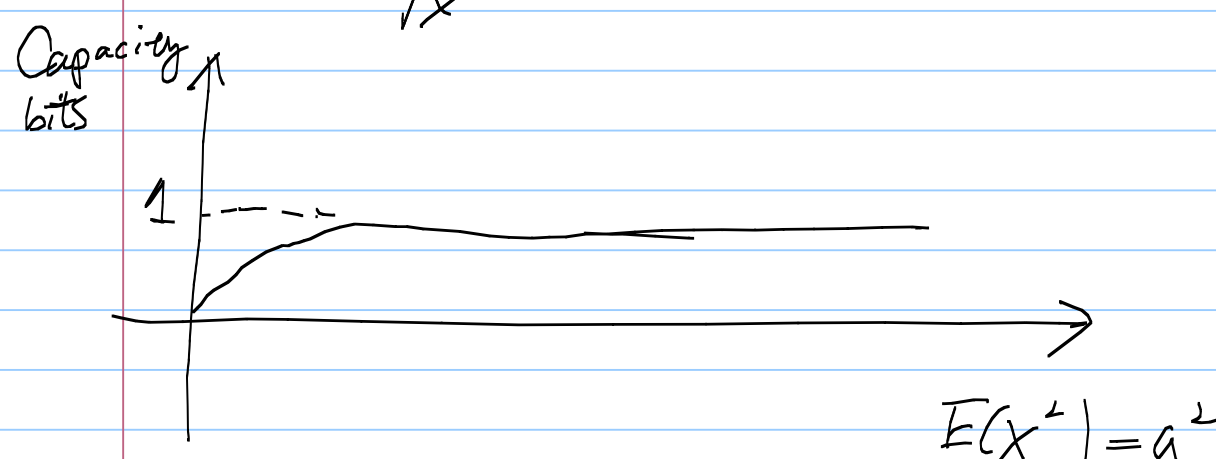
$$= E_{XY}\left(\log \frac{P_{XY}(X,Y)}{P_X(X)\,P_Y(Y)}\right)$$

$$= \sum_{x=0,1} \int_{y=-\infty}^{\infty} P(X=x)\frac{1}{\sqrt{2\pi}}\, e^{-\frac{(y-x)^2}{2\times 1}}$$

$$\cdot \log\left(\frac{P(X=x)\frac{1}{\sqrt{2\pi}}\, e^{-\frac{(y-x)^2}{2}}}{P(X=x)\left(P(X=a)\frac{1}{\sqrt{2\pi}}{}^{-\frac{(y-a)^2}{2}} + P(X=-a)\frac{1}{\sqrt{2\pi}}\, e^{-\frac{(y+a)^2}{2}}\right)}\right) dy$$

This can be evaluated numerically for any p values.

Step 3: $\max\limits_{P_X} I(X;Y)$



The second example

Find out the channel capacity of an i.i.d Gussian channel $N_i \sim Gsn(0,1)$ with the "Power constraint γ". Here X_i does not neet to be of antipodal values but only need to satisfy $E(X^2) \le \gamma$.

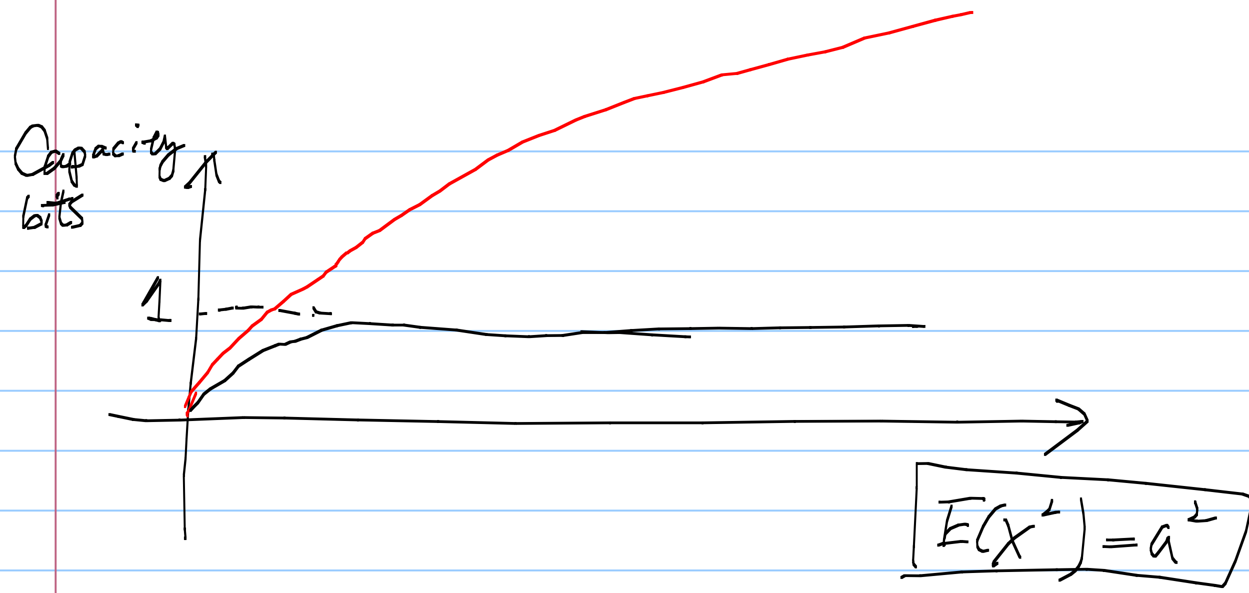
Ans: Step 1: Choose any P_X satisfying
$$E(X^2) \le \gamma$$

Step 2: Find $I(X;Y)$

Step 3: $\max\limits_{P_X : E(X^2) \le \gamma} I(X;Y)$

With power constraint: $p^* \sim Gsn(0, \gamma)$

$$\text{Capacity} = I_{P_X^*}(X;Y) \quad \text{(bits per symbol usage)}$$
$$= \frac{1}{2} \log(1+\gamma) \leftarrow \text{(exercise)}$$



We can also impose different constraints on P_X, ex: $P(|X| > M) = 0$