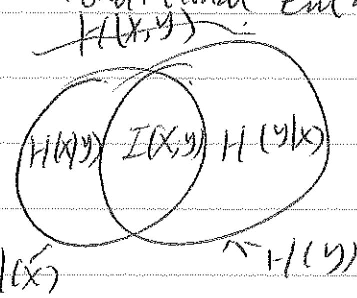
2/12/14 Class 9

review Conditional entropy, mutual entropy



BSC, entropy example

9.8 Channel coding Theorem

Discrete memoryless source → Channel encoder → Discrete memoryless channel → Channel decoder → Destination

noise → Discrete memoryless channel

Channel coding Theorem

i) let discrete memoryless source with alphabet Φ and entropy $H(\Phi)$ with symbol ~~rate~~ per T_s seconds.

let a discrete memoryless channel have capacity C and used per T_c ~~seconds~~ seconds.

If $\frac{H(\Phi)}{T_s} \le \frac{C}{T_c}$

but does not show how → There exists scheme to achieve zero error probability

ii) if $\frac{H(\Phi)}{T_s} > \frac{C}{T_s}$ 50% error.

9.9 Differential Entropy and mutual Information for Continuous ensembles

a continuous random variable X with probability density $f_X(x)$

differential entropy $h(X) = \int_{-\infty}^{\infty} f_X(x) \log_2 \left[\frac{1}{f_X(x)}\right] dx$

ex 9.7 Uniform Distribution $f_X(X) = \begin{cases} \frac{1}{a} & 0 < x < a \\ 0 & \text{otherwise} \end{cases}$

$h(X) = \int_0^a \frac{1}{a} \log(a)\, dx = \log a$

Ex 9.8 Gaussian Distribution

For arbitrary two random variable X and Y

$$\int_{-\infty}^{\infty} f_Y(x) \log_2 \left(\frac{f_X(x)}{f_Y(x)}\right) dx \le 0$$

$$h(Y) \le -\int_{-\infty}^{\infty} f_Y(x) \log_2 f_X(x)\, dx$$

i) suppose X and Y have the same mean μ and variance σ^2

ii) $X \sim N(\mu, \sigma^2)$ $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$

$$h(Y) \le -\log_2 e \int_{-\infty}^{\infty} f_Y(x)\left(-\frac{(x-\mu)^2}{2\sigma^2} - \log(\sqrt{2\pi}\sigma)\right) dx$$

use the facts $\int_{-\infty}^{\infty} f_Y(x)\, dx = 1$ $\int_{-\infty}^{\infty} (x-\mu)^2 f_Y(x)\, dx = \sigma^2$

we have $h(Y) \le h(X) = \frac{1}{2}\log_2(2\pi e \sigma^2)$ — 1

in othe word

i) the entropy is not function of mean

ii) Gaussian has the largest entropy if equal ~~energy~~ variance with any other distribution.

mutual Information, continuous r.v. X Y.

$$I(X;Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) \log \left(\frac{f_X(x|y)}{f_X(x)}\right) dx\, dy$$

Properties 1. $I(X;Y) = I(Y;X)$

2. $I(X;Y) \ge 0$

3. $I(X;Y) = h(X) - h(X|Y) = h(Y) - h(Y|X)$

where $h(X|Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) \log_2 \left[\frac{1}{f_X(x|y)}\right] dx\, dy$

9.10 Information Capacity Theorem

band-limited, power-limited, Gaussian channel.

Let X_k, $k = 1, 2 \dots K$ denote continuous r.v. obtained by uniform sampling of $X(t)$ at rate $2B$ sample/second these samples are transmitted in T seconds over noisy channel. ~~over noisy cha~~

$K = 2BT$

AWGN (additive white Gaussian noise)

$X_K \to \Sigma \to Y_K$ $\quad Y_K = X_K + N_K, \quad k = 1, 2 \ldots K$

N_K $\quad N_K \sim N(0, \sigma^2) \quad \sigma^2 = N_0 B.$

$E[X_K^2] = P \quad E(X_K) = 0.$

Define Information capacity of the channel

$$C = \max_{f_{X_K}(x)} \{ I(X_K; Y_K) : E\{X_K^2\} = P \}$$

$$I(X_K; Y_K) = h(Y_K) - h(Y_K | X_K) = h(Y_K) - h(N_K)$$

to maximize $I(X_K, Y_K)$, $h(N_K)$ independent

X_K has to be Gaussian so that $h(Y_K)$ is also Gaussian (Gaussian + Gaussian = Gaussian) with max entropy (under the same power)

$C = I(X; Y_K)$: X_K Gaussian, $E(X_K^2) = P$.

1) $h(Y_K) = \frac{1}{2} \log_2 [2\pi e (P + \sigma^2)]$

from (1) $\quad Y_K$ received power $P + \sigma^2$

2 $\quad h(N_K) = \frac{1}{2} \log_2 (2\pi e \sigma^2)$ $\quad$ (2)

3 $\quad C = \frac{1}{2} \log_2 (1 + \frac{P}{\sigma^2})$ bits per transmission

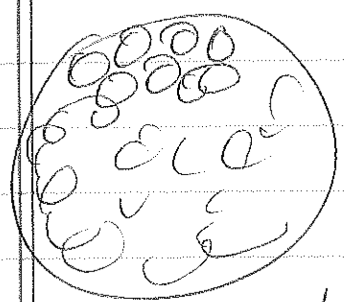
remember $K = 2BT$

$C = B \log_2 (1 + \frac{P}{N_0 B})$ bits per second.

Trade off between bandwidth and power

with same C $\begin{cases} B \uparrow & P \downarrow \\ B \downarrow & P \uparrow \end{cases}$

Sphere packing: received vector of n bits is Gaussian distributed with mean equal to the tx code word and variance $n\sigma^2$, where σ^2 is noise power with high probability, the received vector lies in a sphere of radius $\sqrt{n\sigma^2}$, centered on the tx code word. The sphere is contained in a larger of radius $\sqrt{n(P^2+\sigma^2)}$, where $n(P+\sigma^2)$ is the average rx power



The volume of sphere of rx vector $A_n[n(P+\sigma^2)]^{n/2}$

the volume of the decoding sphere $A_n(n\sigma^2)^{n/2}$

$$\underbrace{\frac{A_n[n(P+\sigma^2)]^{n/2}}{A_n(n\sigma^2)^{n/2}}}_{\text{the maximum number of non overlapping}} = \left(1+\frac{P}{\sigma^2}\right)^{n/2} = 2^{(n/2)\log_2(1+P/\sigma^2)}$$

Take $\log_2$ is ②

Ex 9.9 Reconfiguration of constellation for reduced power

Fig. 9.15

9.11 Implication of the information Capacity Theorem

$P = E_b C$, where E_b is tx power per bit.

$$\frac{C}{B} = \log_2\left(1+\frac{E_b}{N_0}\frac{C}{B}\right)$$

$$\frac{E_b}{N_0} = \frac{2^{C/B}-1}{C/B}$$

① 1. $\left(\frac{E_b}{N_0}\right)_\infty = \lim_{B\to\infty}\left(\frac{E_b}{N_0}\right) = \log 2 = 0.693$

$$C_\infty = \lim_{B\to\infty} C = \frac{P}{N_0}\log_2 e$$

2. capacity boundary

③ trade off of Bandwidth v.s Power