Discrete Memoryless Channel

$$X\left\{\begin{matrix} x_0 \\ x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_{J-1} \end{matrix}\right. \quad X \rightarrow P(y_k \mid x_j) \rightarrow Y \quad \left.\begin{matrix} y_0 \\ y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_{K-1} \end{matrix}\right\} Y$$

- Definition of DMC
  - ⇨ Channel with input X & output Y which is noisy version of X.
  - ⇨ Discrete when both of alphabets X & Y finite sizes.
  - ⇨ Memoryless when no dependency between input symbols.

Discrete Memoryless Channel (cont')

- Channel Matrix (Transition Probability Matrix)

$$P = \begin{bmatrix} p(y_0 | x_0) & p(y_1 | x_0) \ldots & p(y_{K-1} | x_0) \\ p(y_0 | x_1) & \ldots\ldots & p(y_{K-1} | x_1) \\ . & & \\ . & & \\ p(y_0 | x_{J-1}) & \ldots\ldots\ \ldots & p(y_{K-1} | x_{J-1}) \end{bmatrix}$$

➡ The size is J by K

$$\sum_{k=0}^{K-1} p(y_k | x_j) = 1 \text{ for all j}$$

➡ a priori prob. is :

$$P_k = p(x_j), \; j = 0, 1, .., J-1$$

Discrete Memoryless Channel (cont')

- Given a priori prob. $p(x_j)$, and the channel matrix, P then we can find the prob. of the various output symbols, $p(y_k)$ as
  - ⇨ the joint prob. dist'n of X and Y

$$p(x_j, y_k) = p(X = x_j, Y = y_k) = p(Y = y_k / X = x_j) p(X = x_j)$$

$$= p(y_k / x_j) p(x_j)$$

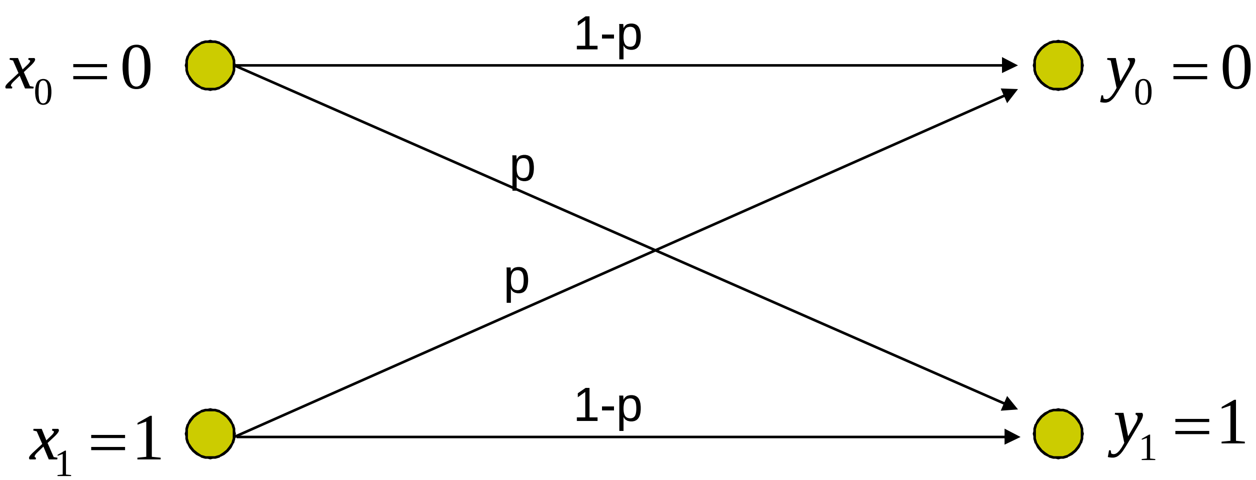
  - ⇨ the marginal prob. dist'n of the output Y,

$$p(y_k) = p(Y = y_k) = \sum_{j=0}^{J-1} p(Y = y_k / X = x_j) p(X = x_j)$$

$$= \sum_{j=0}^{J-1} p(y_k / x_j) p(x_j), \quad for \ \ k = 0, 1, .., K-1$$

Discrete Memoryless Channel(cont')

- **BSC (Binary Symmetric Channel)**



Mutual Information

- **Conditional Entropy**

$$H(X \mid Y = y_k) = \sum_{j=0}^{J-1} p(x_j \mid y_k) \log_2 [\frac{1}{p(x_j \mid y_k)}]$$

- **The mean value**

$$H(X \mid Y) = \sum_{k=0}^{K-1} H(X \mid Y = y_k) p(y_k)$$

$$= \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(x_j, y_k) \log_2 [\frac{1}{p(x_j \mid y_k)}]$$

-> **H(X|Y)** : **a conditional entropy (equivocation)**

The amount of uncertainty remaining about the channel input data after the channel output has been observed.

- **Mutual Information : The uncertainty of the input resolved by observing output**

***I(X;Y) ≡ H(X) - H(X|Y)* ,and** $I(X;Y) = \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 [\frac{p(y_k \mid x_j)}{p(y_k)}]$

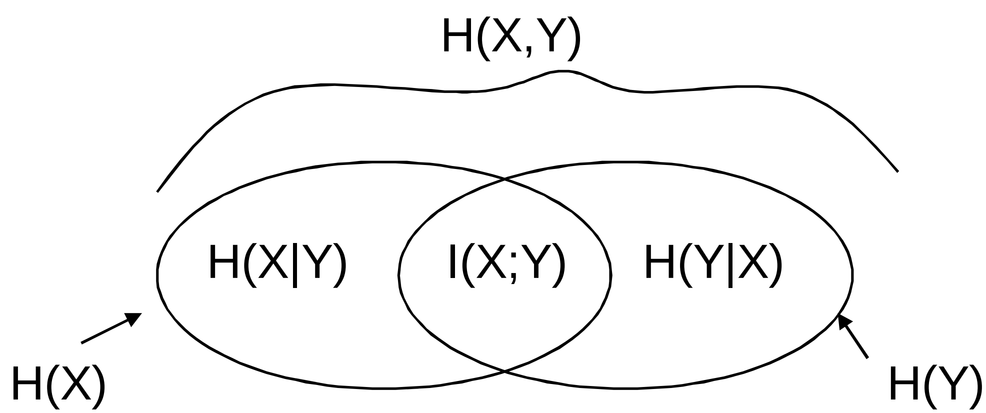
Properties of Mutual Information

(simple ex. needed for 2 by 2 DMC)

- **Symmetric :** $I(X;Y) = I(Y;X)$
- **Non-negative** : $I(X;Y) \geq 0$
- $I(X;Y) = H(Y) - H(Y \mid X)$
- $I(X;Y) = H(X) + H(Y) - H(X,Y)$

where

$$H(X,Y) = \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 [\frac{1}{p(x_j, y_k)}]$$



Channel Capacity

- **For a dms with input X, output Y, &** $p(y_k \mid x_j)$,

$$I(X;Y) = \sum_{j=0}^{J-1}\sum_{k=0}^{K-1} p(x_j, y_k)\log_2\left[\frac{p(y_k \mid x_j)}{p(y_k)}\right]$$

where $p(x_j, y_k) = p(y_k \mid x_j)p(x_j)$, $p(y_k) = \sum_{j=0}^{J-1} p(y_k \mid x_j)p(x_j)$

$\Rightarrow$ **I(X;Y) just depends upon** $\{p(x_j), j = 0,1,2,\ldots, J-1\}$ **, & channel. Since** $\{p(x_j)\}$ **is indep. of the channel, it is possible to maximize I(X;Y) w.r.t.** $\{p(x_j)\}$**.**

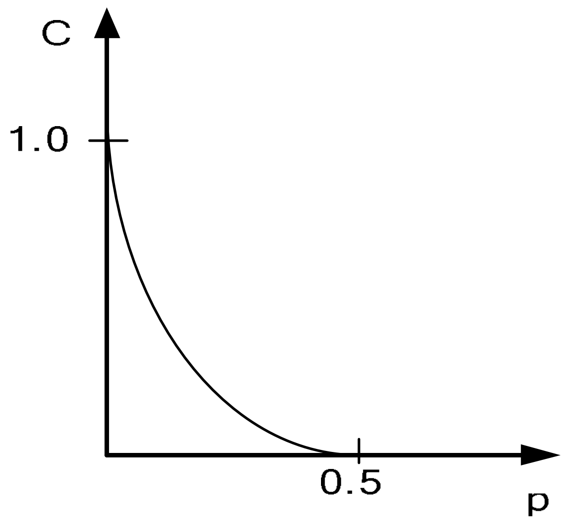
- **Def. of Channel Capacity.**

$$C = \max_{\{p(x_j)\}} I(X;Y) \quad \text{(bits per channel use)}$$

Ex.) for BSC

$$C = \max I(X;Y) = I(X;Y)\Big|_{p(x_0)=0.5}$$

$$\therefore C = 1 + p\log_2 p + (1-p)\log_2(1-p) = 1 - H(p)$$



Channel Coding Theorem

- For reliable communication , needs channel encoding & decoding.
  "any coding scheme which gives the error as small as possible, and which is efficient enough that code rate is not too small?"

=> Shannon's second theorem (noisy coding theorem)

Let dms with alphabet X have entropy H(X) and produce symbols once every Ts, and dmc have capacity C and be used once every Tc . Then,

i) if $\frac{H(X)}{T_s} \leq \frac{C}{T_c}$, there exists a coding scheme.

ii) if $\frac{H(X)}{T_s} > \frac{C}{T_c}$, it is not possible to transmit with arbitrary small error.

Ex.) for BSC with $p_0 = 0.5$

The condition for reliable comm. ,

$$\frac{1}{T_s} \le \frac{C}{T_c}$$

Let $\frac{T_c}{T_s}$ be r , then $r \le C$

$\therefore$ for $r \le C$, there exists a code (with code rate less than or equal to C) capable of achieving an arbitrary low probability of error.

“ The code rate $r = \frac{k}{n}$ where k is k-bit input, and n is n-bit coded bits,.”

Differential Entropy

- Differential Entropy

$$h(X) = \int_{-\infty}^{\infty} f_x(x) \log_2 [\frac{1}{f_x(x)}] dx$$

where $f_x(x)$ is p.d.f.

-extension into continuous r.v.

- Basis to derive the Channel Capacity Theorem

Maximum Differential Entropy for Specified Variance

Find p.d.f. for which $h(x)$ is maximum, subject to

$$i)\quad \int_{-\infty}^{\infty} f_X(x)dx = 1 \quad , \quad ii)\quad \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)dx = \sigma^2 = const$$

where μ is the mean, and σ^2 is the variance

Since σ^2 is a measure of average power, it is to find maximization with constraint of constant power

Maximum Differential Entropy for Specified Variance

Sol. is based on calculus of variation & use of Lagrange multiplier

$$I = \int_{-\infty}^{\infty} \left[-f_X(x)\log_2 f_X(x) + \lambda_1 f_X(x) + \lambda_2 (x-\mu)^2 f_X(x) \right] dx$$

$\rightarrow$ I should be stationary to get maximum entropy.

- The desired form of $f_X(x)$ is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-u)^2}{2\sigma^2} \right)$$

$\therefore$ Gaussian p.d.f.

- The maximum entropy

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

$*$ Gaussian channel model is so widely utilized.

Mutual Information for Continuous r.v

$$I(X:Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\log_2\left[\frac{f_X(x\,|\,y)}{f_X(x)}\right]dxdy$$

where $f_{X,Y}(x,y)$ is the joint *pdf* of X&Y, &

$f_X(x\,|\,y)$ is the conditional *pdf* of X, given that $Y = y$

$$\Rightarrow \underset{f_{x_k}(x)}{\text{Max}}\{I(X_k:Y_k) : \text{E}\left[X_k^2\right] = P]$$

leads to the most important channel capacity

Shannon's Channel Capacity Theorem

For bandlimited, power limited Gaussian channels

$$C = B\log_2\left(1+\frac{P}{N}\right) \text{ (bits/s)}$$

The capacity of a channel of bandwidth B, perturbed by additive white gaussian noise of psd $N_0/2$, and limited in bandwidth to B, P is the average transmitted power, and N is the noise (N_oB)

- It is not possible to transmit at rate higher than C reliability by any means.
- It does not say how to find coding and modulation to achieve maximum capacity, but it indicates that approaching this limit, the transmitted signal should have statistical property approximately to Gaussian noise.

Bandwidth efficient diagram

Define ideal system as $R_b = C$,

$P = E_b C$ where E_b is the Tx energy per bit

Then, $\frac{C}{B} = \log_2(1 + \frac{E_b C}{N_o B})$

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

- For infinite bandwidth channel

$$\left(\frac{E_b}{N_o}\right)_{B->\infty} = \ln 2 = 0.693 = -1.6dB$$

$$C_\infty = \lim_{B->\infty} C = \frac{P}{N_o} \log_2 e$$