情報理論12

藤田 一寿

津山工業高等専門学校情報工学科 講師
電気通信大学先進理工学科 協力研究員

ver.20150701

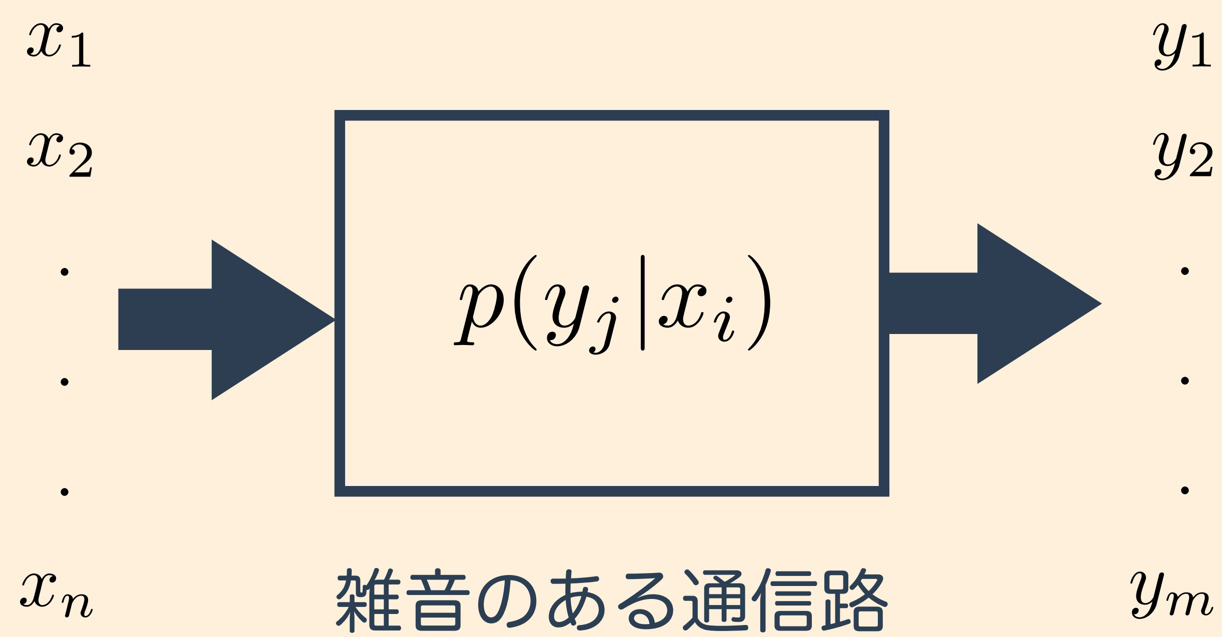
雑音のある通信路

雑音の妨害での情報伝送量

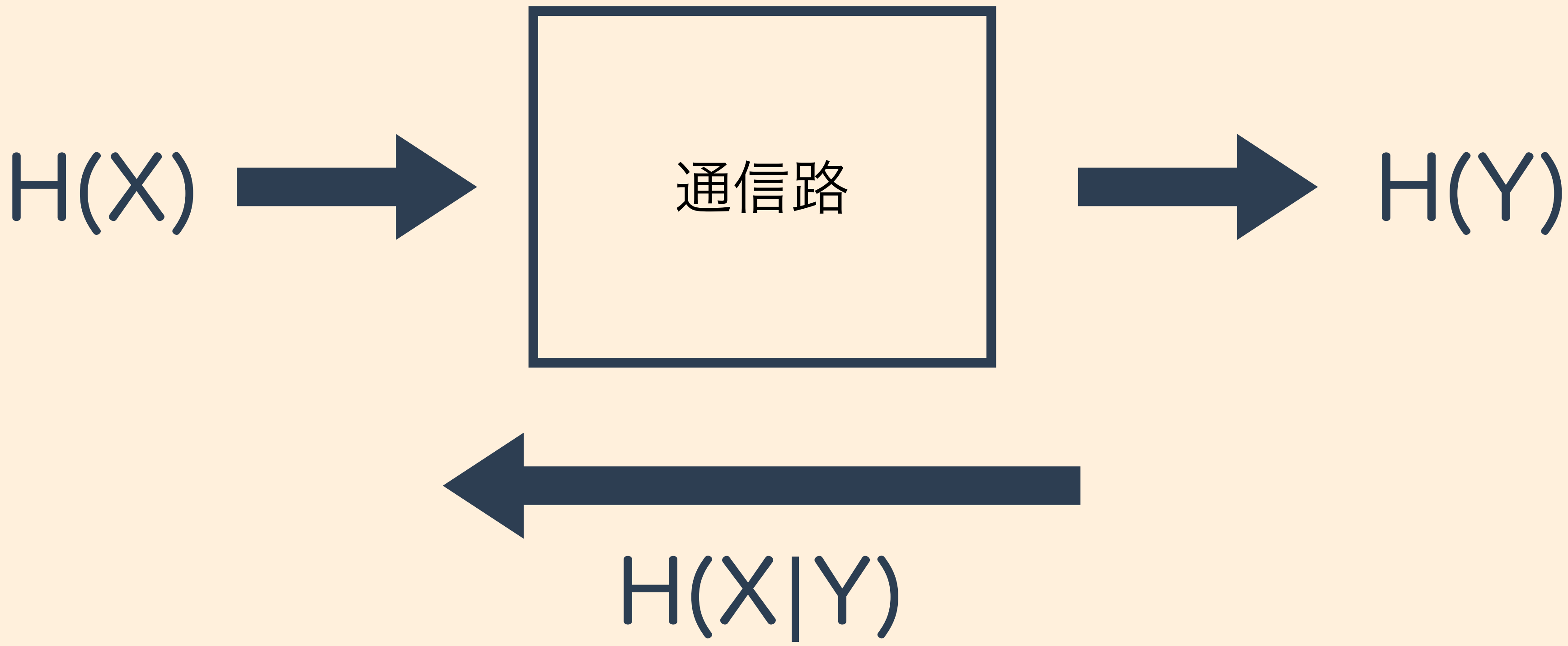
通信する上で雑音がない場合はほぼ無い

どのようにして確実に情報を送るか

雑音のある通信路



- 入力信号x1を送った時，なにが受信されるか
  - y1かも知れないしy2かも知れない
  - 雑音（ノイズ）により変わる
- 確率的に取り扱うと，yjが出る確率はp(yj|xi)と表すことができる．



- 信号源は1秒辺りH(X)のエントロピーを持つ.
- 信号系列x1, x2, x3,…は雑音の影響によりy1, y2, y3,…に変換される.
- 受信信号は1秒辺りH(Y)のエントロピーを持つとする.
- 受信される信号の確率p(yj|xi)が分かれば，受信した信号からなんとなく元の信号が分かる.
- 受信信号が分かるという条件のもとでの情報源のエントロピーはH(X|Y)と書ける．情報源が持つ曖昧さを表す.

情報伝達速度

受信信号により得られた情報量Rを情報伝達速度と呼ぶ.

$$R = H(X) - H(X|Y)$$

情報源のエントロピー　受信信号の曖昧さ

情報伝送速度Rは，式の通り送信信号と受信信号の相互情報量になっている.

$$R = H(X) - H(X|Y) = I(X,Y)$$

X，Y入れ替えても同じ

少し具体的に計算してみる

送信信号のエントロピー

$$H(X) = -\sum_i p(x_i)\log p(x_i)$$

受信信号が生成される確率

$$p(y_i) = \sum_i p(x_i)p(y_j|x_i)$$ xiとyjの同時確率を求め周辺化している

受信信号のエントロピー

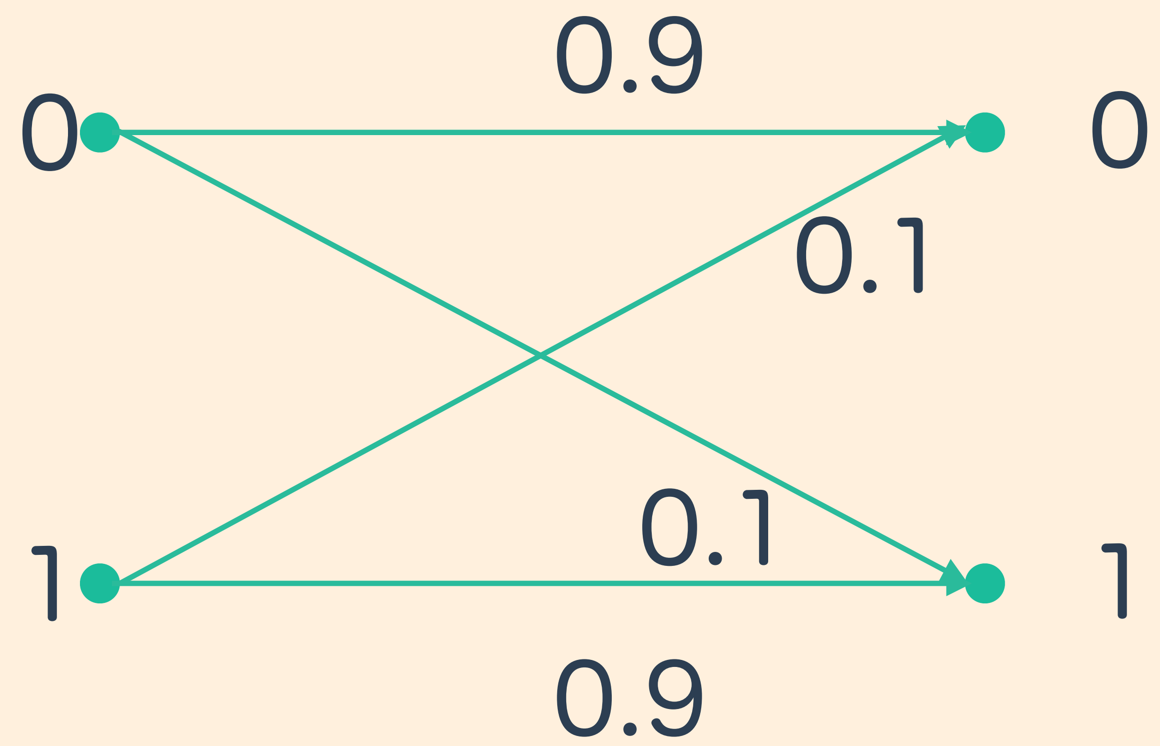
$$H(Y) = -\sum_j p(y_j)\log p(y_j)$$

受信信号が分かっているものとして，送信信号の曖昧さ

$$p(x_i|y_j) = \frac{p(y_j|x_i)p(x_i)}{p(y_j)}$$

$$H(X|Y) = -\sum_i \sum_j p(y_j)p(x_i|y_j)\log p(x_i|y_j)$$

例



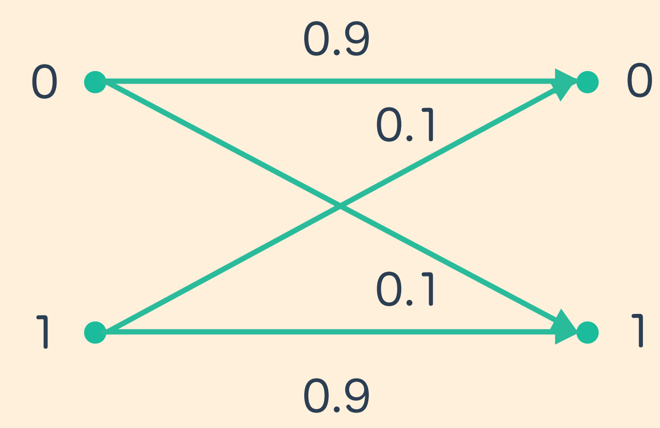
雑音のため，0.1の確率で信号が変わってしまう通信路

$$p(y=0|x=0)=p(y=1|x=1)=0.9$$

$$p(y=0|x=1)=p(y=1|x=0)=0.1$$

0もしくは１を等確率に１秒間に１０００個出力するとする.
信号源のエントロピーは

$$
\begin{aligned}
H(X) &= 1000 \times (-\sum_i p(x_i) \log p(x_i)) \\
&= 1000 \times (-\frac{1}{2} \log \frac{1}{2} - \frac{1}{2} \log \frac{1}{2}) \\
&= 1000 \times (-\log \frac{1}{2}) \\
&= 1000
\end{aligned}
$$



1000桁の2進数なので当たり前.
信号が伝送されると雑音により各文字が0.1の確率で変異する.
そう考えると1000個の文字のうち900個が変異せずに到達すると考えられる．と考えると900ビットの情報が伝送されたといえるのだろうか.

曖昧さ

$$
\begin{aligned}
H(X|Y) =& 1000 \times (-\sum_i \sum_j p(y_j)p(x_i|y_j)\log p(x_i|y_j)) \\
=& 1000 \times (-p(y_0)p(x_0|y_0)\log p(x_0|y_0) - p(y_0)p(x_1|y_0)\log p(x_1|y_0) \\
& - p(y_1)p(x_0|y_1)\log p(x_0|y_1) - p(y_1)p(x_1|y_1)\log p(x_1|y_1)) \\
=& 1000 \times (-0.5 \times 0.9\log 0.9 - 0.5 \times 0.1\log 0.1 \\
& - 0.5 \times 0.9\log 0.9 - 0.5 \times 0.1\log 0.1 \\
=& 100 \times (- \times 0.9\log 0.9 - \times 0.1\log 0.1) \\
=& 469
\end{aligned}
$$

情報伝送速度Rは

$$
\begin{aligned}
R &= 1000 - 469 \\
&= 531
\end{aligned}
$$

900個の文字があたっていても，どれがあたっているかわからないので0.1の確率で雑音が入ってしまっても情報は半分に減る.

ノイズによりでたらめな文字列になったら

- 文字列がランダムに受信している事になり，各文字の出現確率が0.5となる．結局曖昧さはH(X|Y)=1000ビットとなる.

- つまり，最大転送速度は0ビットとなり，何も情報を伝えていないことになる.

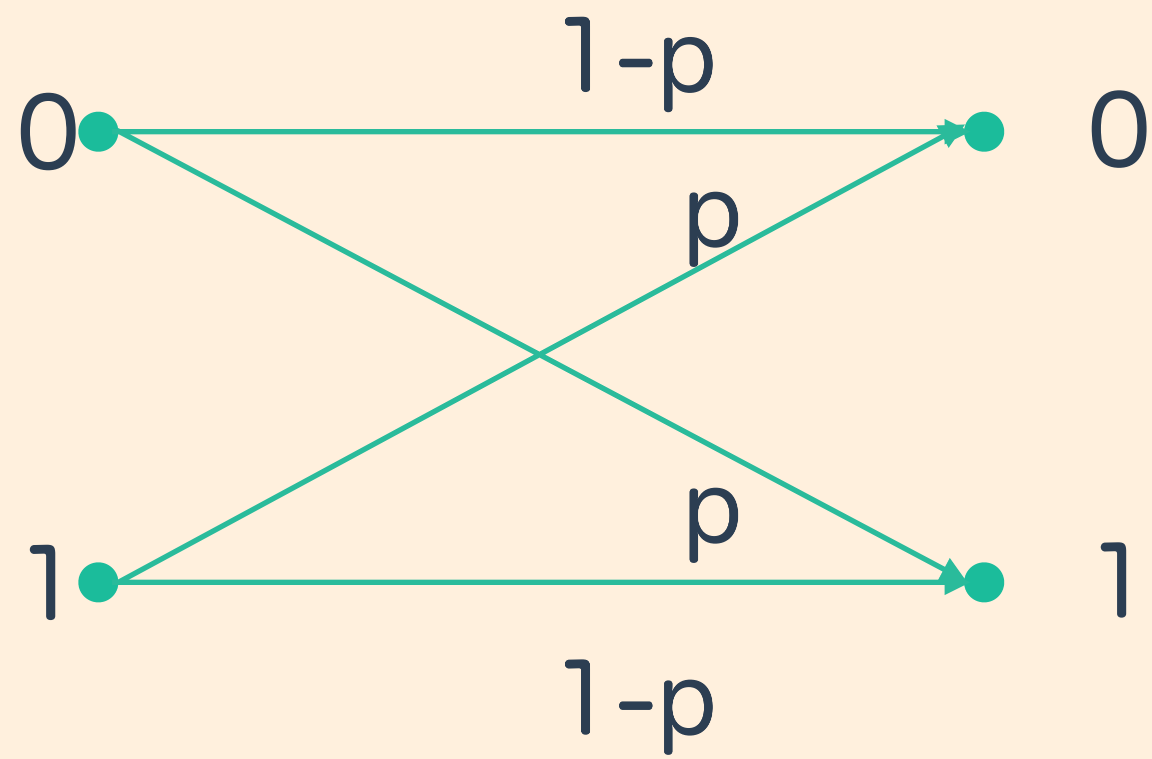
通信路容量

- 通信路容量を大きくするには

$$C = \max_{p_i} R$$

- 雑音の影響が少ない信号の確率を増やす.

2元対称通信路



通信路行列

$$P = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

$$P = \begin{bmatrix} p(0|0) & p(1|0) \\ p(0|1) & p(1|1) \end{bmatrix}$$

$$H(X) = -p_0 \log p_0 - p_1 \log p_1$$

$$H(Y) = -q_0 \log q_0 - q_1 \log q_1$$

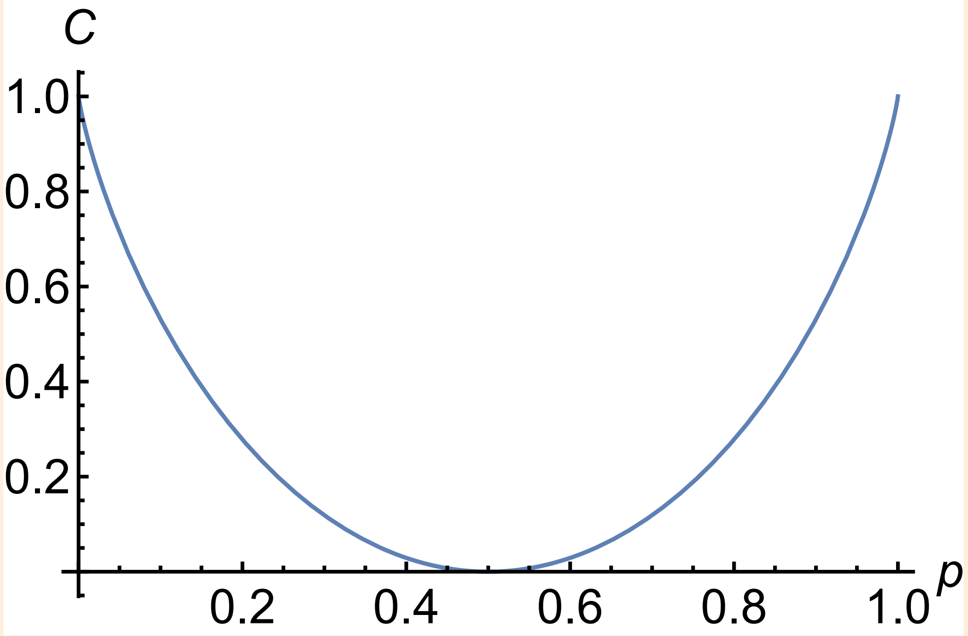
$$H(Y|X) = -p \log p - (1-p) \log(1-p)$$

$$\begin{aligned} R = I(X,Y) &= H(Y) - H(Y|X) \\ &= -q_0 \log q_0 - q_1 \log q_1 + p \log p + (1-p) \log(1-p) \end{aligned}$$

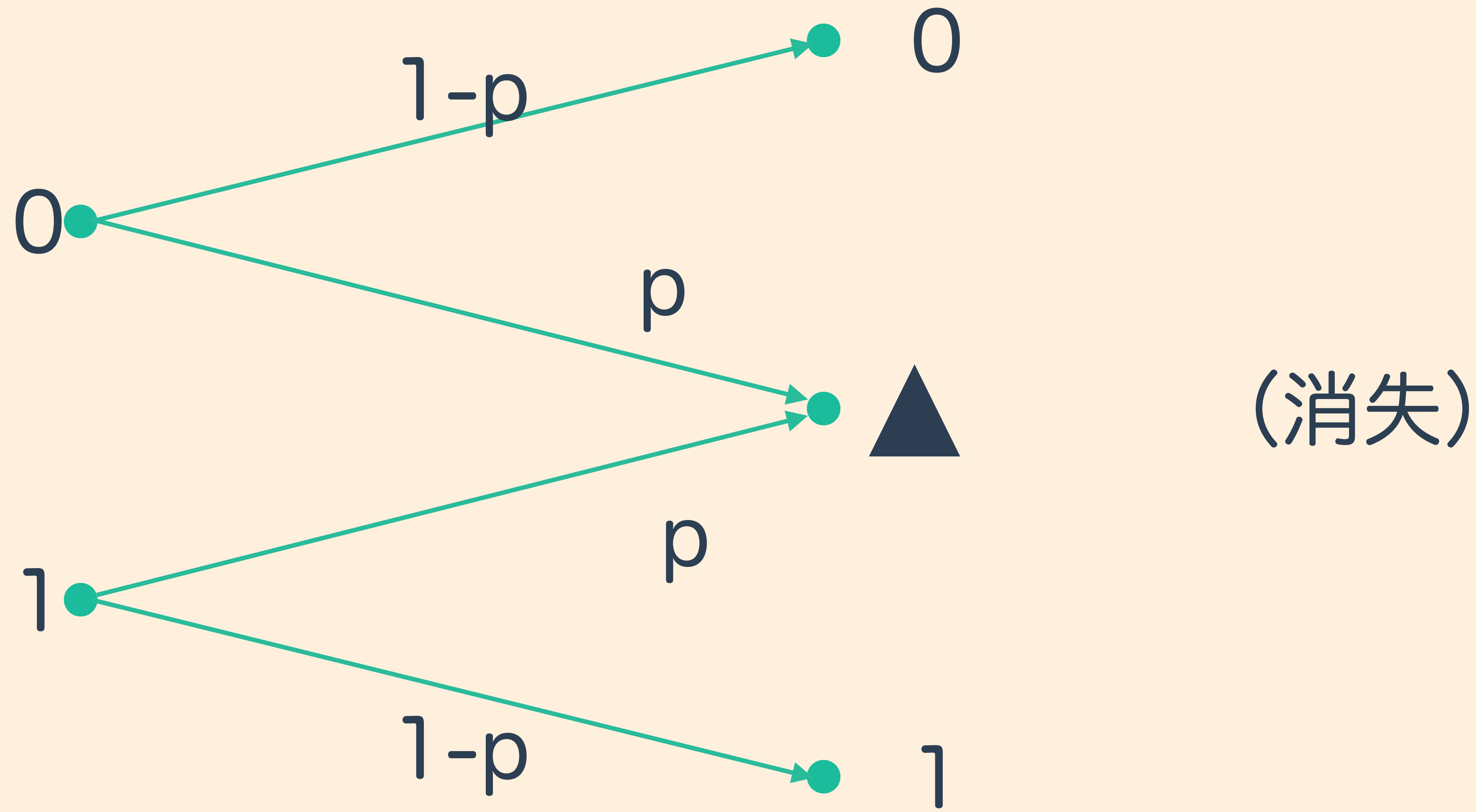
Rが最大になるのはq0=q1=0.5 (p0=p1=0.5)ときき．よって

$$C = 1 + p \log p + (1-p) \log(1-p)$$

p=0.5の時，伝送速度は0となる．p=1の場合は常に信号が入れ替わることが分かっているので，確実に信号が伝わる事となり，伝送速度が1となる．

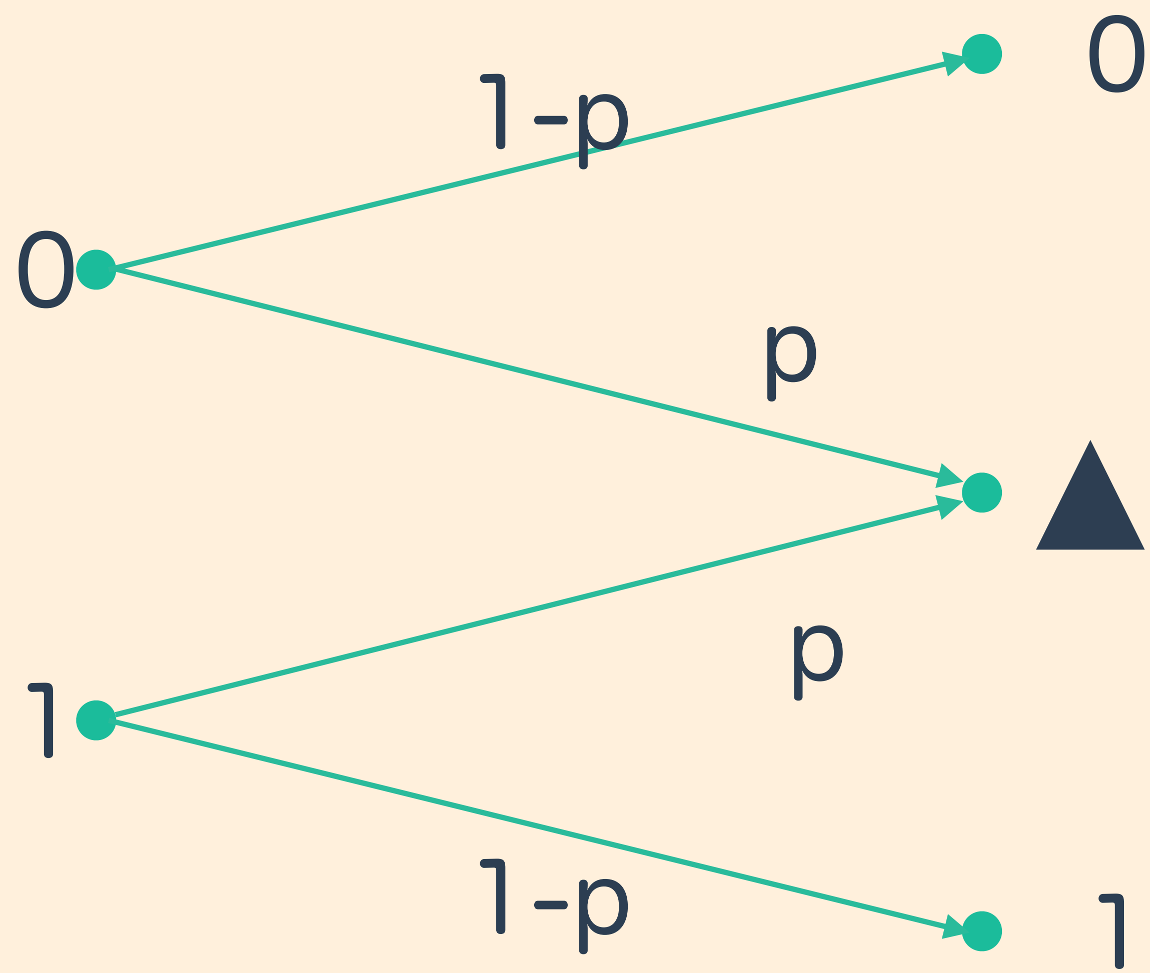


2元消失通信路



0，1を1秒ごとに1つ送る通信路
出力信号は0，1，△
信号は確率pで△に変わる

$$P = \begin{bmatrix} 1-p & p & 0 \\ 0 & p & 1-p \end{bmatrix}$$



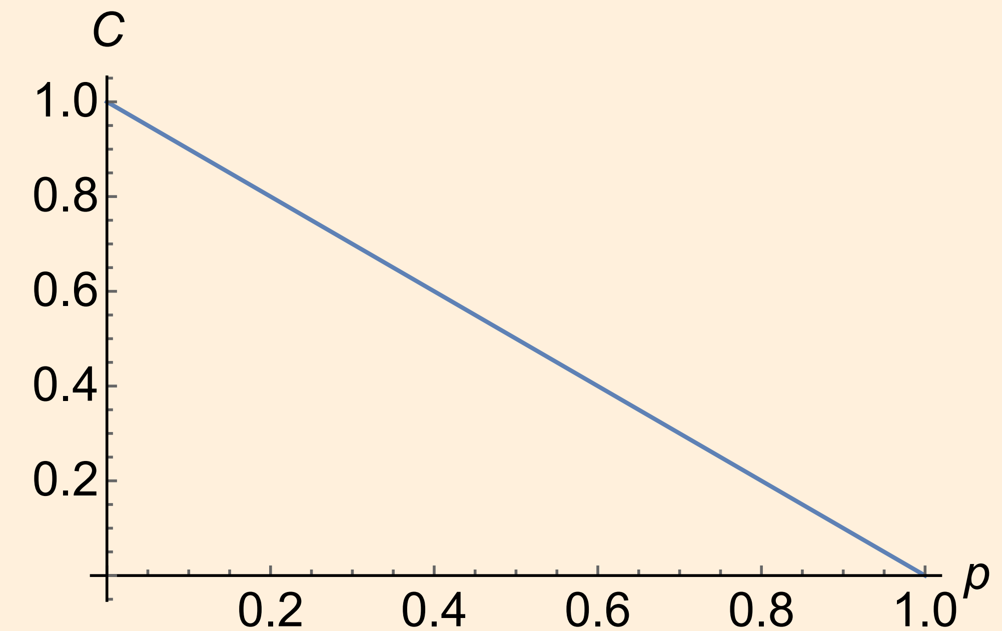
$$H(Y) = -p_0(1-p)\log(p_0(1-p)) - p_1(1-p)\log(p_1(1-p))$$
$$= -(1-p)(p_0\log p_0 + p_1\log p_1 - (1-p)\log(1-p))$$

$$H(Y|X) = \sum_i \sum_j p(x_i)p(y_j|x_i)\log p(y_j|x_i)$$
$$= -(1-p)\log(1-p)$$

$$R = H(Y) - H(Y|X) = (1-p)H(X)$$

Rを最大化するp0，p1はそれぞれ0.5であるから，最大通信路容量Cは

$$C = 1 - p$$



2元対称通信路とくらべてみる.

p = 0.1の場合，例２ではC = 0.9となり，例１の0.53と比べ大きい.