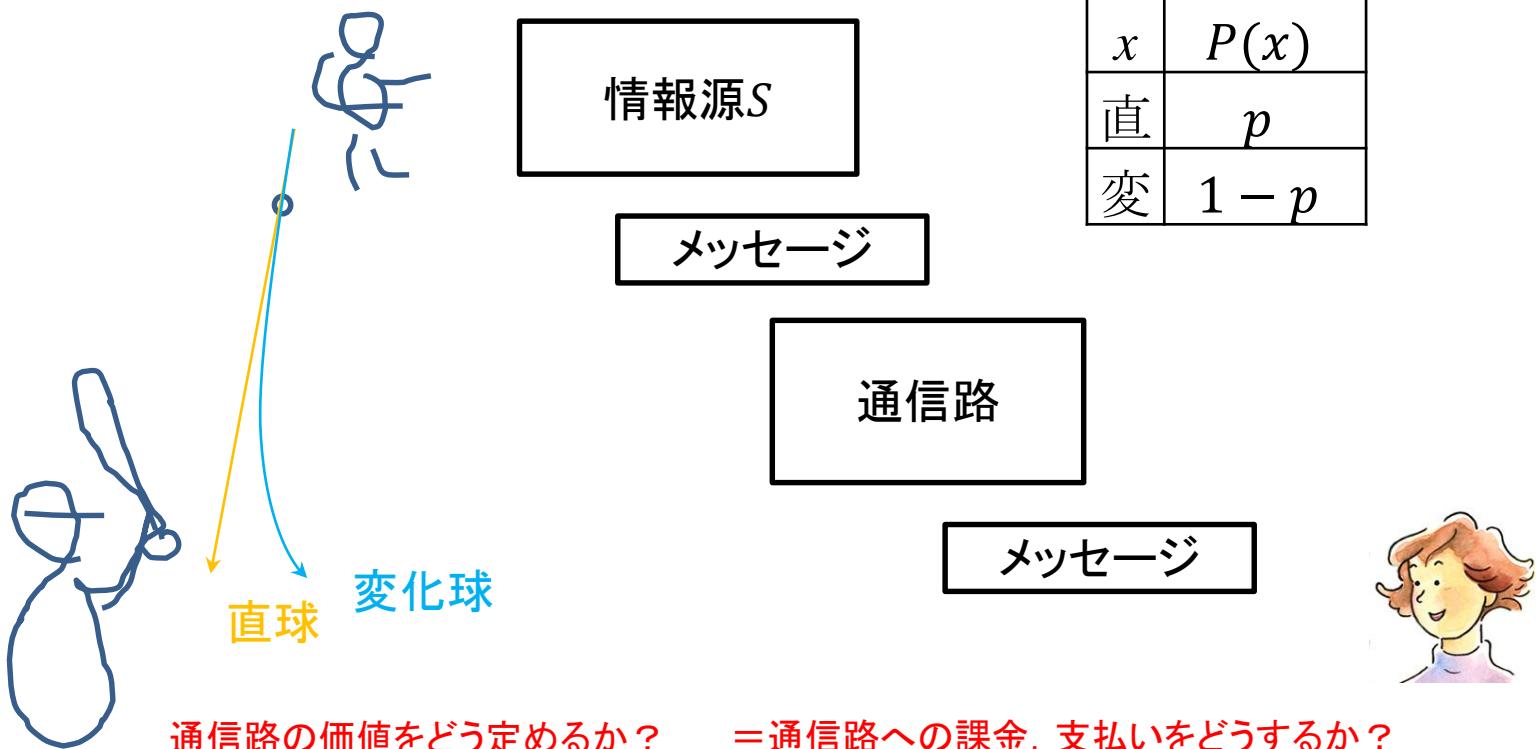


相互情報量と通信路モデル

例題 一 野球中継



通信路の価値をどう定めるか？ 一通信路への課金、支払いをどうするか？

メッセージの価値をどれだけ保全するか

メッセージの価値をどう定めるか？

情報源Sの不確かさをどれだけ減らしたか？

$$S \text{ のエントロピー } H_1(S) = -(p \log_2 p + (1 - p) \log_2(1 - p)) \quad \dots \text{ 平均符号長の下限}$$

メッセージの平均的価値の大きさ = メッセージの平均情報量 = メッセージが減らした情報源記号の不確かさ

情報量

出現確率 p の事象 e が起きたことを知らせるメッセージ m のもたらす「サプライズ」($=m$ の情報量)

$$\log_2 \frac{1}{p} = -\log_2 p \text{ ビット}$$

(例) 変化球の確率が $\frac{1}{4}$ であるとき、実際にそうであるということを伝えるメッセージの情報量：

$$\log_2 \frac{1}{\frac{1}{4}} = -\log_2 4 = 2 \text{ ビット}$$

情報量

2つの事象 $\{e_1, e_2\}$ のいずれかが起きるとする.

e_1 の生起確率を p_1 , e_2 の生起確率を p_2 とする.

- e_1 が起きたことを知らせるメッセージの情報量:

$$\log_2 \frac{1}{p_1} = -\log_2 p_1 \text{ ビット}$$

- e_2 が起きたことを知らせるメッセージの情報量:

$$\log_2 \frac{1}{p_2} = -\log_2 p_2 \text{ ビット}$$

- メッセージの平均的な情報量:

$$p_1 \log_2 \frac{1}{p_1} + p_2 \log_2 \frac{1}{p_2} = -(p_1 \log_2 p_1 + p_2 \log_2 p_2) \text{ ビット}$$

情報量

n 個の事象 $\{e_1, \dots, e_n\}$ が与えられたとき、 e_i の生起確率を p_i とすると、 e_i の起きたことを知らせるメッセージの平均的な情報量：

$$p_1 \log_2 \frac{1}{p_1} + \cdots + p_n \log_2 \frac{1}{p_n} = - \sum_{i=1}^n p_i \log_2 p_i \text{ ビット}$$

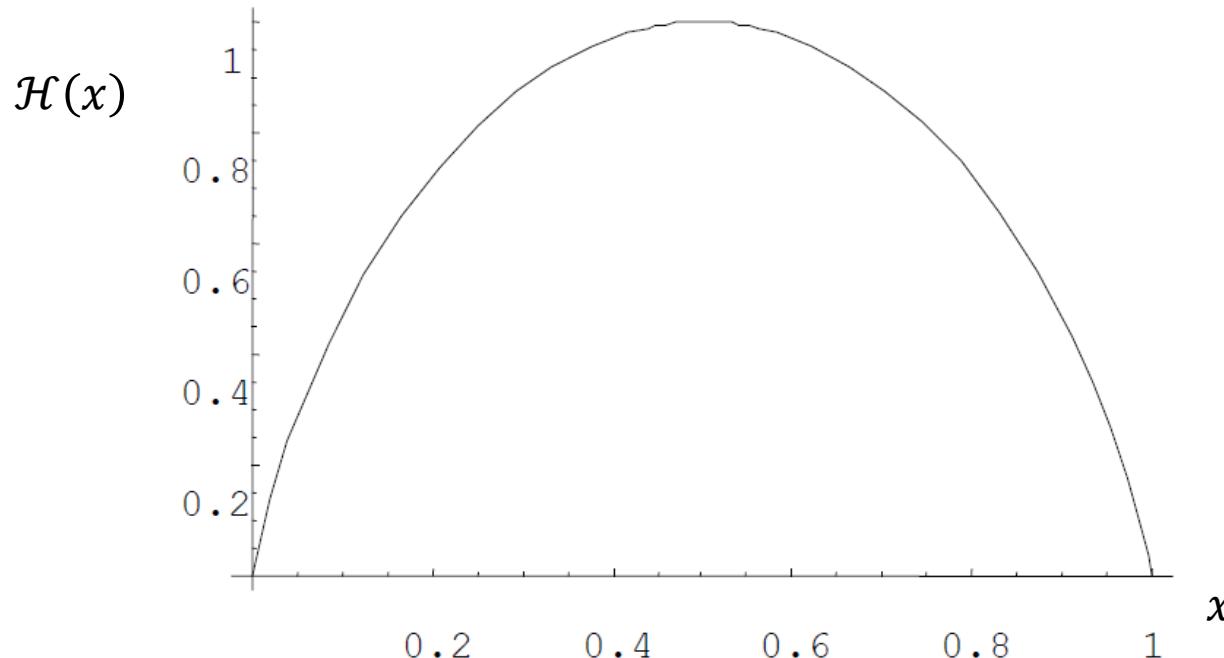
- 例1:{直: 1/2, 変: 1/2}
… 直でも、変でも1ビットの情報が得られる。平均情報量は1ビット。
- 例2:{直: 1/4, 変: 3/4} … $1/4$ の確率で $\log_2 4 = 2$ ビットの情報が得られるが、残り $3/4$ の確率で得られる情報量は $\log_2 \frac{4}{3} \approx 0.415$ ビットだけである。平均すれば、

$$-\left(\frac{1}{4} \log_2 \frac{1}{\frac{1}{4}} + \frac{3}{4} \log_2 \frac{1}{\frac{3}{4}}\right) \approx \left(\frac{1}{4} \times 2 + \frac{3}{4} \times 0.415\right) \approx 0.811 \text{ ビット}$$

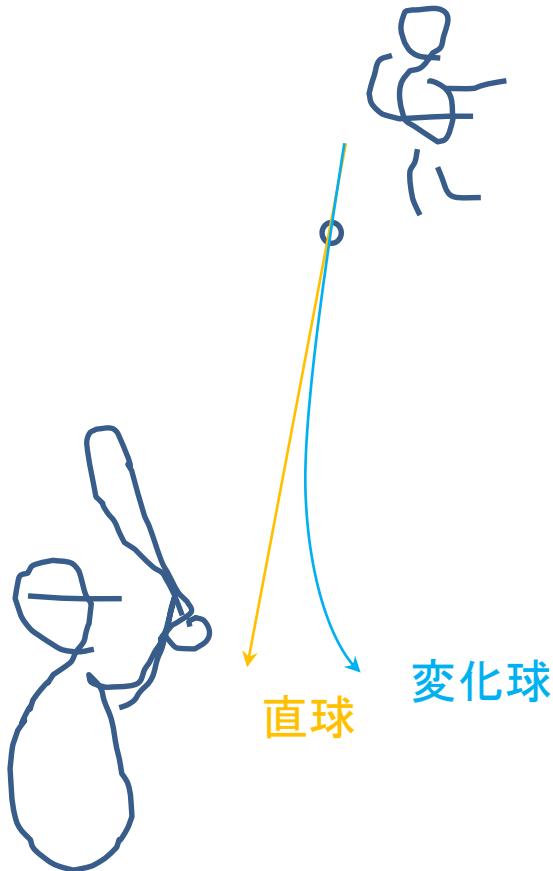
情報量

- 情報量: メッセージが不確かさを減らす度合いを定量化
- エントロピー関数: 不確かさの度合いの表現に用いる.

$$\mathcal{H}(x) \equiv x \log_2 \frac{1}{x} + (1 - x) \log_2 \frac{1}{1 - x}$$

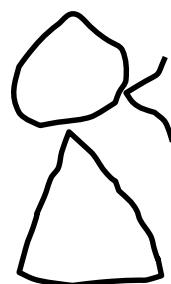


知りたいことのヒントにしかならないメッセージ



	直	変
前	0.4	0.2
横	0.1	0.3

{前, 横}



{直, 変}?



検査

検査の効用評価

	病気	健康
陽性	$\frac{2}{1000}$	$\frac{10}{1000}$
陰性	$\frac{8}{1000}$	$\frac{920}{1000}$

$$P(\text{病氣}) = \frac{10}{1000} = 1\%$$

$$P(\text{病氣}|\text{陽性}) = \frac{P(\text{陽性}, \text{病氣})}{P(\text{陽性})} = \frac{\frac{2}{1000}}{\frac{2}{1000} + \frac{10}{1000}} = \frac{1}{6} \approx 16.7\%$$

$$\text{偽陽性率} = P(\text{陽性}|\text{健康}) = \frac{P(\text{陽性, 健康})}{P(\text{健康})} = \frac{\frac{10}{1000}}{\frac{10}{1000} + \frac{920}{1000}} = \frac{1}{93} \approx 1.1\%$$

$$\text{偽陰性率} = P(\text{陰性} | \text{病氣}) = \frac{P(\text{陰性}, \text{病氣})}{P(\text{病氣})} = \frac{\frac{8}{1000}}{\frac{2}{1000} + \frac{8}{1000}} = 80\%$$

$$(陽性)尤度比 = \frac{1 - 偽陰性率}{偽陽性率} = \frac{P(\text{陽性}|\text{病気})}{P(\text{陽性}|\text{健康})} = \frac{\frac{2}{1000}}{\frac{2}{1000} + \frac{8}{1000}} = \frac{\frac{2}{10}}{\frac{1}{1000} + \frac{920}{1000}} = \frac{\frac{2}{10}}{\frac{922}{1000}} = \frac{2}{93} = \frac{\frac{P(\text{病気}|\text{陽性})}{P(\text{健康}|\text{陽性})}}{\frac{P(\text{病気})}{P(\text{健康})}} = \frac{\text{陽性の事後オッズ}}{\text{陽性の事前オッズ}} = 18.6$$

相互情報量

- 何も知らないときは、「直」は1ビットの情報量、「変」も1ビットの情報量をもつから、直／変メッセージは平均1ビットの情報量をもつことになる。
- 前／横メッセージを得たとしよう。
 - 「前」だったとすれば、
 - 「直」は、 $\log_2 \frac{3}{2} \approx 0.585$ ビットの情報をもたらす。
 - 「変」は、 $\log_2 3 \approx 1.585$ ビットの情報をもたらす。
 - 平均すると、 $\frac{2}{3} \log_2 \frac{3}{2} + \frac{1}{3} \log_2 3 = \log_2 3 - \frac{2}{3} \approx 0.918$ ビットの情報がもたらされる。
 - 「横」だったとすれば、
 - 直／変メッセージの平均情報量は $\frac{1}{4} \log_2 4 + \frac{3}{4} \log_2 \frac{1}{3} = 2 - \frac{3}{4} \log_2 3 \approx 0.811$
- 前／横メッセージを得たときの、直／変メッセージの平均情報量は
 - $\frac{3}{5} \times \left(\log_2 3 - \frac{2}{3} \right) + \frac{2}{5} \times \left(2 - \frac{3}{4} \log_2 3 \right) = \frac{2}{5} + \frac{3}{10} \log_2 3 \approx 0.8755$ ビット
- 従って、前／横メッセージの平均的な貢献(=平均情報量)は、 $1 - \left(\frac{2}{5} + \frac{3}{10} \log_2 3 \right) = \frac{3}{5} - \frac{3}{10} \log_2 3 \approx 0.1245$ ビット

相互情報量

確率変数 X のエントロピー(平均的な不確かさ)を $H(X)$ とすると,

$$H(\text{直変}) = \mathcal{H}\left(\frac{1}{2}\right) = 1$$

$$H(\text{直変|前}) = \mathcal{H}\left(\frac{1}{3}\right) \approx 0.91830$$

$$H(\text{直変|横}) = \mathcal{H}\left(\frac{1}{4}\right) \approx 0.81128$$

$$H(\text{直変|前横}) = \frac{3}{5} \mathcal{H}\left(\frac{1}{3}\right) + \frac{2}{5} \mathcal{H}\left(\frac{1}{4}\right) = \frac{2}{5} + \frac{3}{10} \log_2 3 \approx 0.87549$$

$$H(\text{直変}) - H(\text{直変|前横}) = \frac{3}{5} - \frac{3}{10} \log_2 3 \approx 0.1245$$

一方,

$$H(\text{前横}) = \mathcal{H}\left(\frac{3}{5}\right) \approx 0.97095$$

$$H(\text{前横|直}) = \mathcal{H}\left(\frac{4}{5}\right) \approx 0.72193$$

$$H(\text{前横|変}) = \mathcal{H}\left(\frac{2}{5}\right) \approx 0.97095$$

$$H(\text{前横|直変}) = \frac{1}{2} \mathcal{H}\left(\frac{4}{5}\right) + \frac{1}{2} \mathcal{H}\left(\frac{2}{5}\right) \approx 0.84644$$

$$H(\text{前横}) - H(\text{前横|直変}) = \frac{3}{5} - \frac{3}{10} \log_2 3 \approx 0.1245$$

このように,

$$H(\text{直変}) - H(\text{直変|前横}) = H(\text{前横}) - H(\text{前横|直変}) = \frac{3}{5} - \frac{3}{10} \log_2 3 \approx 0.1245$$

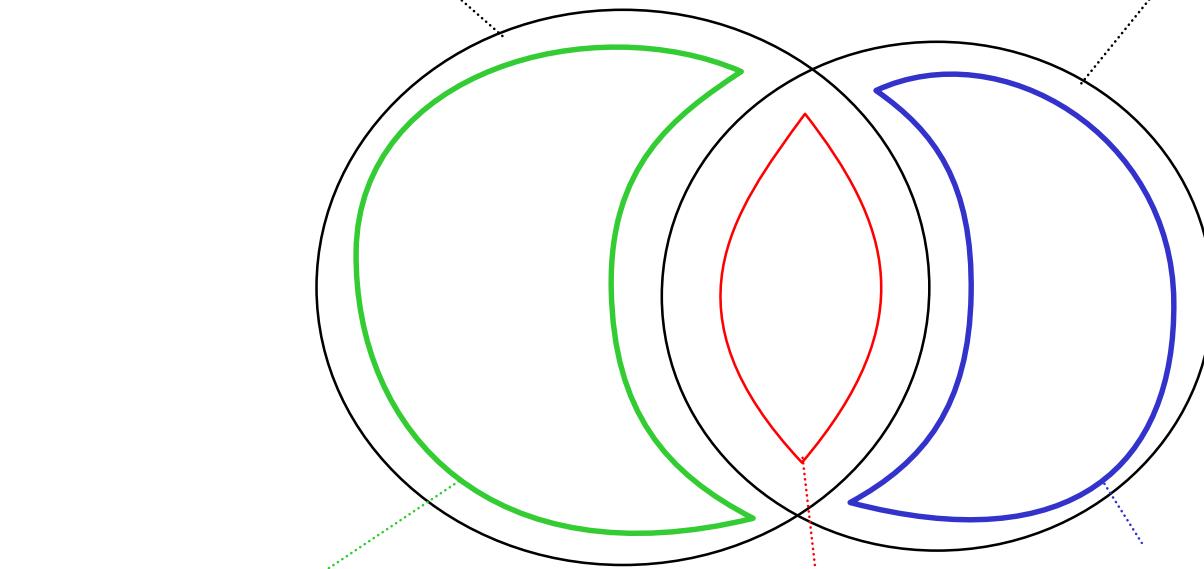
相互情報量の概念

$$H(\text{直変})=1$$

直／変の不確かさ

$$H(\text{前横}) \approx 0.97095$$

前／横の不確かさ



$$H(\text{直変|前横}) \approx 0.87549$$

前／横がわかったときの直／変の不確かさ

$$H(\text{前横|直変}) \approx 0.84644$$

直／変が分かったときの前／横の不確かさ

$$I(\text{直変;前横}) = H(\text{直変}) - H(\text{直変|前横}) = H(\text{前横}) - H(\text{前横|直変}) \approx 0.1245$$

相互情報量

前／横がわかったときに減少する直／変の不確かさ
直／変が分かったときに減少する前／横の不確かさ

相互情報量の概念

不確かなメッセージでも情報を伝達することができる. \Rightarrow 伝達される情報量の定量化

(例)上でみたように、次の球種が直球か変化球か知りたいとき、投手が前を見ていたか、横を見ていたかを教えてもらうことができれば、何らかのヒントが得られる(ただし、確率モデルがわかっていると仮定できる時).

$P(x, y)$		Y		$P(x)$
		直	変	
X	前	0.4	0.2	0.5
	横	0.1	0.3	0.5
$P(y)$		0.6	0.4	

- 前／横によって減少する不確かさの平均:

$$H(Y) - H(Y|X) \approx 0.1245$$

- X と Y の相互情報量:

$$I(X; Y) = H(Y) - H(Y|X)$$

相互情報量の性質

- X と Y の結合エントロピー:

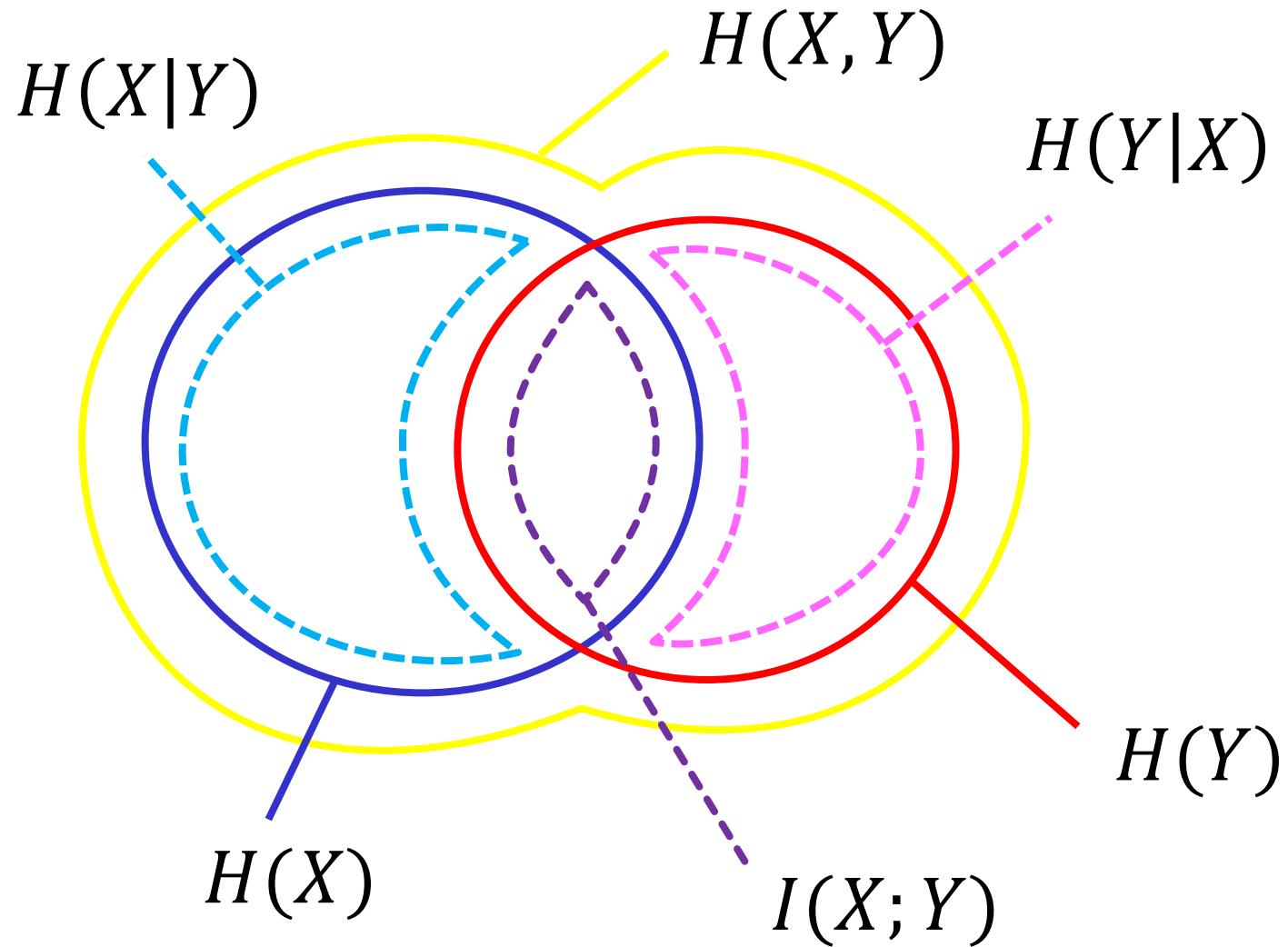
$$H(x, y) = - \sum_x \sum_y P(x, y) \log_2 P(x, y)$$

- 次の関係が成り立つ.

$$\begin{aligned} I(X; Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \\ &= H(X) + H(Y) - H(X, Y) \end{aligned}$$

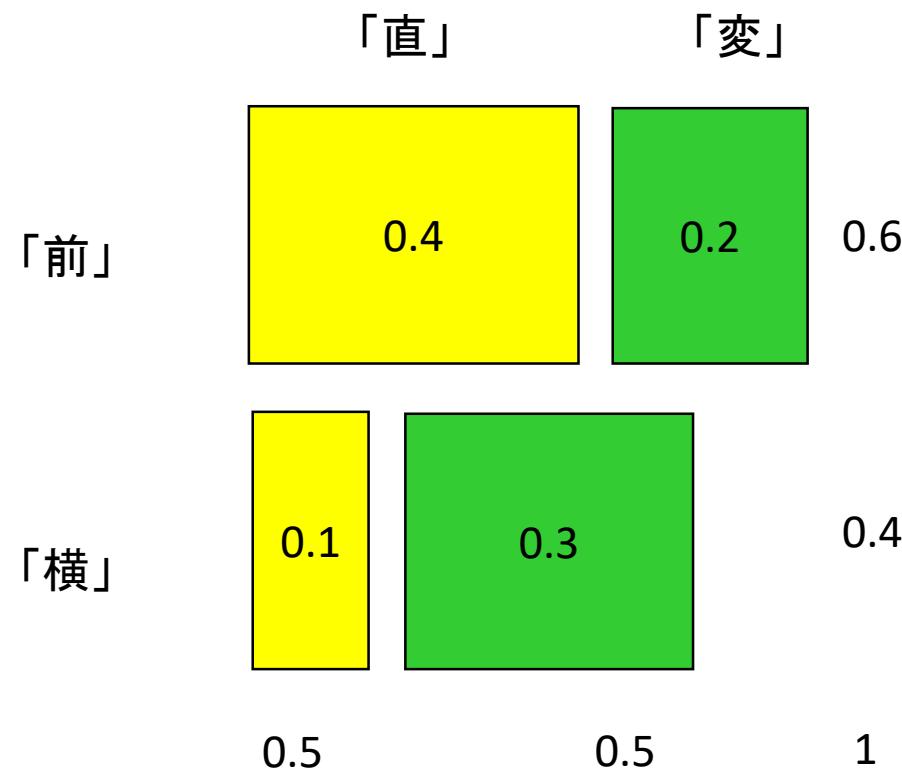
$$0 \leq I(X; Y) \leq \min\{H(X), H(Y)\}$$

相互情報量の性質



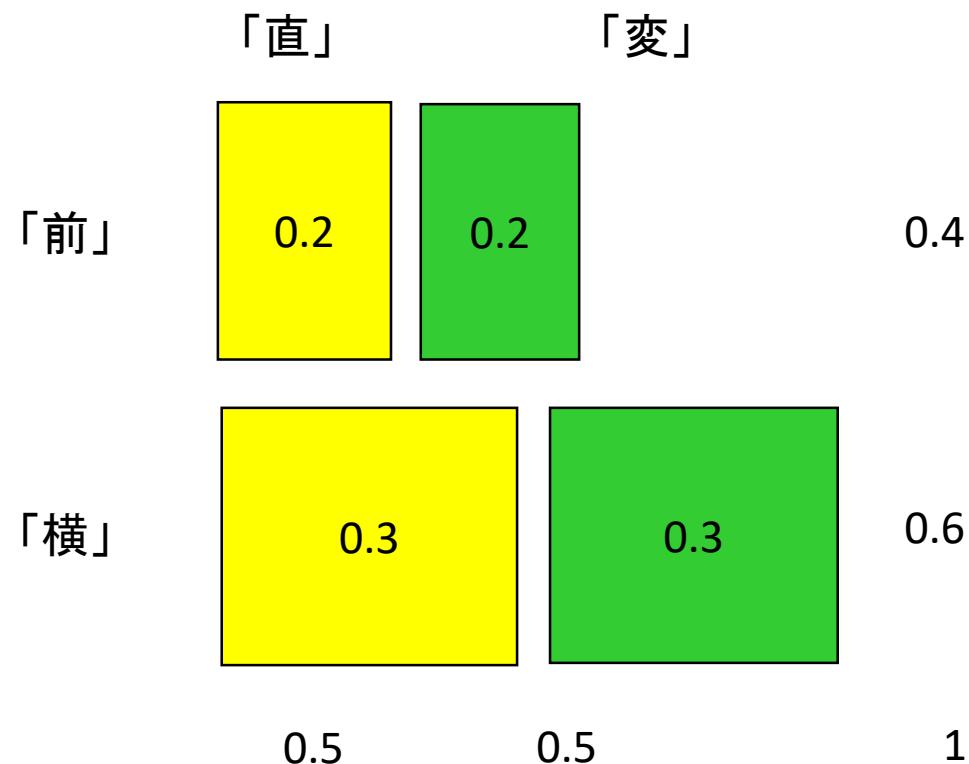
相互情報量の性質

「前／横」から得られる「直／変」に関する情報量

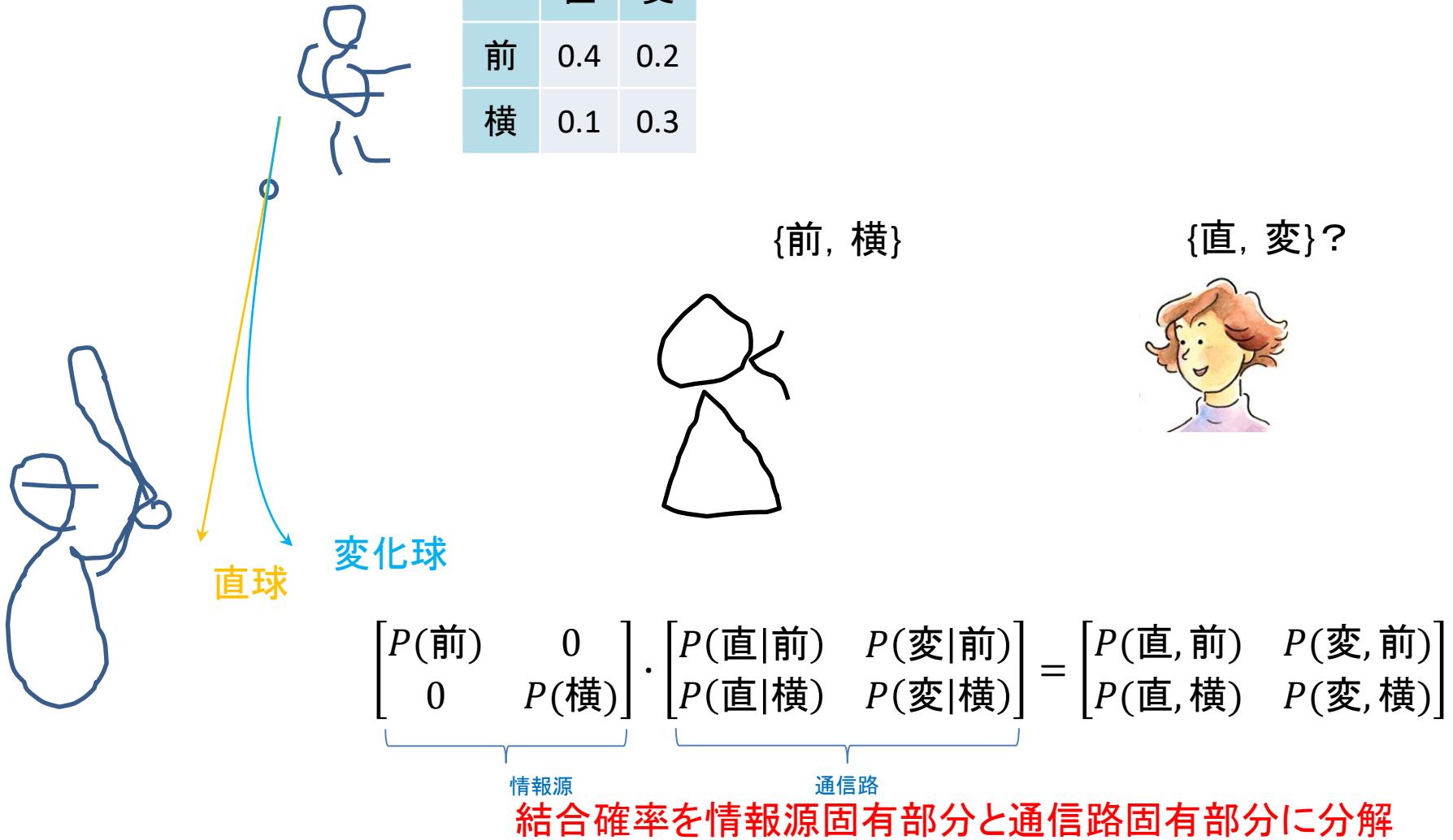


相互情報量の性質

「前／横」から得られる「直／変」に関する情報量 ... もう一つのケース

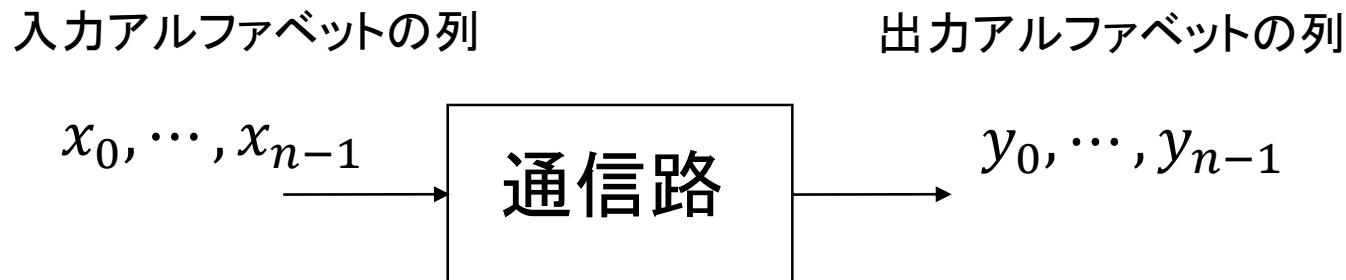


通信路の概念



通信路モデル

通信路を条件付確率で表し、統計的性質を調べる。



$$P_{Y_0, \dots, Y_{n-1} | X_0, \dots, X_{n-1}}(y_0, \dots, y_{n-1} | x_0, \dots, x_{n-1})$$

x_0, \dots, x_{n-1} が与えられたとき y_0, \dots, y_{n-1} が出力される確率

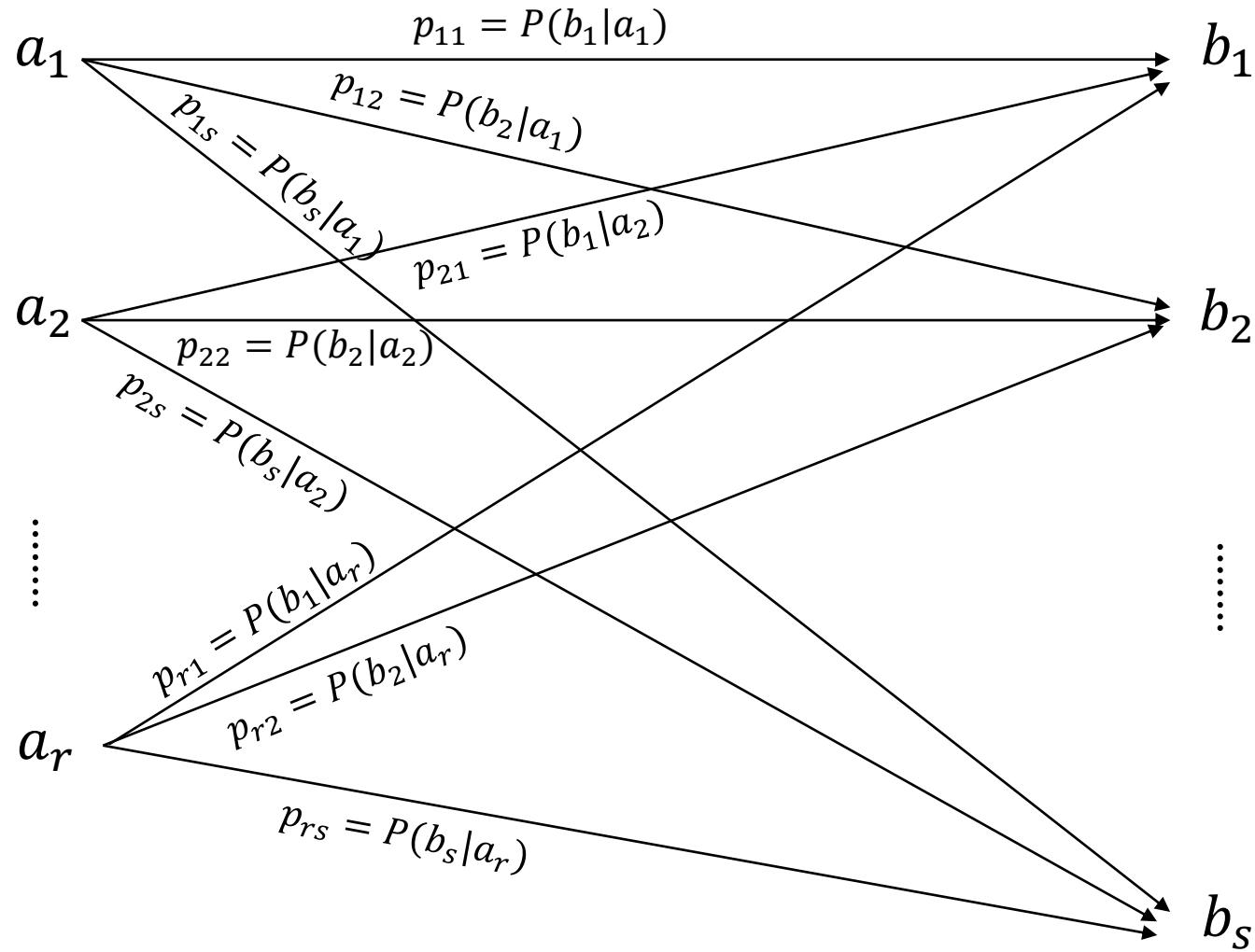
通信路モデル

- 記憶のない通信路: 各時点の出力の現れ方がその時点の入力には関係するが、それ以外の時点の出力にも入力にも独立である。
- 記憶のない定常通信路: 記憶のない通信路のなかで、時間をずらしても統計的性質が変わらないもの。
- 記憶のない定常通信路について,

$$P_{Y_0, \dots, Y_{n-1} | X_0, \dots, X_{n-1}}(y_0, \dots, y_{n-1} | x_0, \dots, x_{n-1}) = \prod_{i=0}^{n-1} P_{Y_i | X_i}(y_i | x_i)$$

通信路モデル

通信路の表示



通信路行列

- 通信路行列: 記憶のない定常通信路の表現

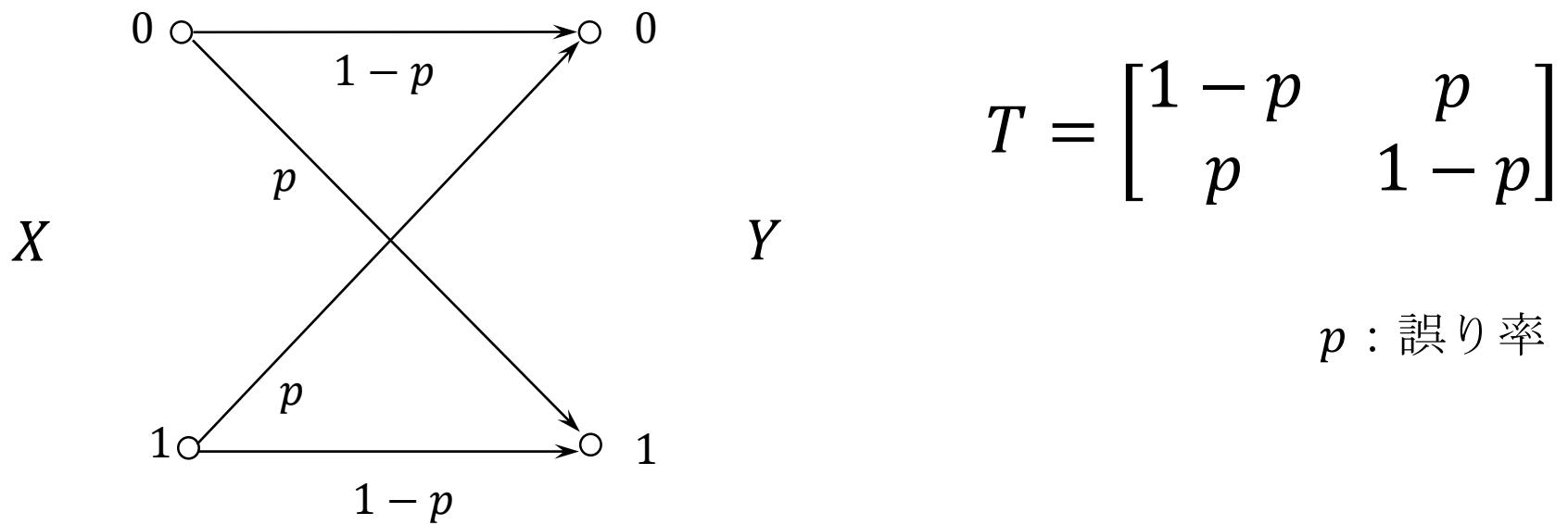
$$T = \begin{bmatrix} p_{11} & \cdots & p_{1s} \\ \vdots & \ddots & \vdots \\ p_{r1} & \cdots & p_{rs} \end{bmatrix}$$

- 入力アルファベットの生起確率: $p = (p_1, p_2, \dots, p_r)$
出力アルファベットの生起確率: $q = (q_1, q_2, \dots, q_s)$ のとき,

$$\begin{aligned} q &= (q_1, q_2, \dots, q_s) \\ &= pT \\ &= (p_1, p_2, \dots, p_r) \begin{bmatrix} p_{11} & \cdots & p_{1s} \\ \vdots & \ddots & \vdots \\ p_{r1} & \cdots & p_{rs} \end{bmatrix} \\ &= (p_1 p_{11} + p_2 p_{21} + \cdots + p_r p_{r1}, \\ &\quad p_1 p_{12} + p_2 p_{22} + \cdots + p_r p_{r2}, \\ &\quad \dots, \\ &\quad p_1 p_{1s} + p_2 p_{2s} + \cdots + p_r p_{rs}) \end{aligned}$$

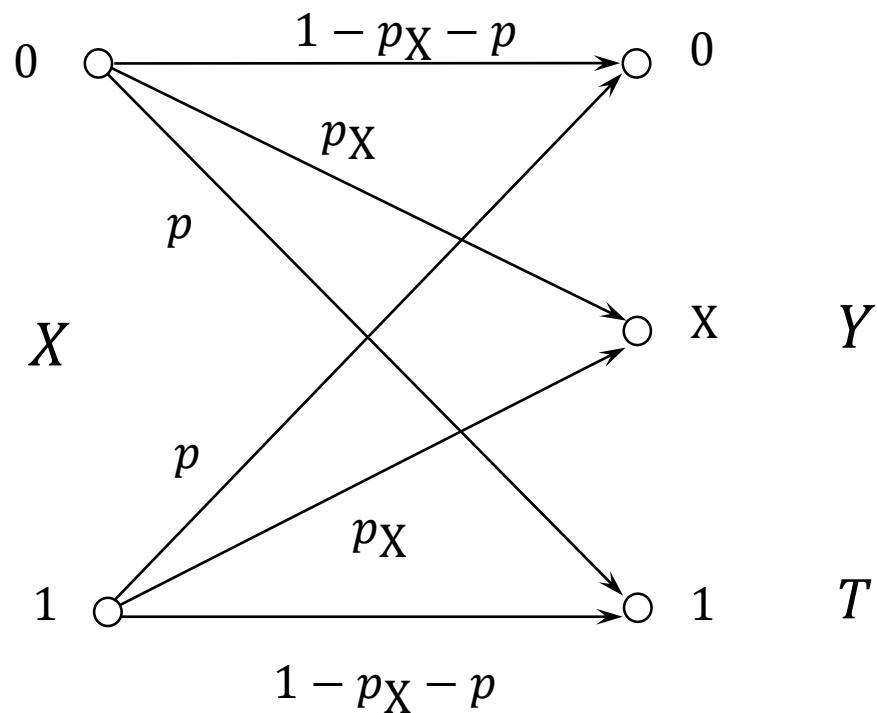
基本的な通信路

2元対称通信路(Binary Symmetric Channel, BSC)



基本的な通信路

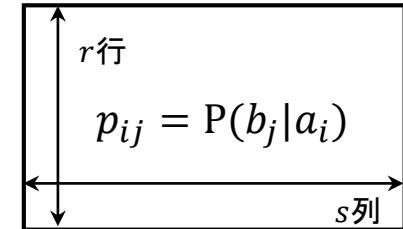
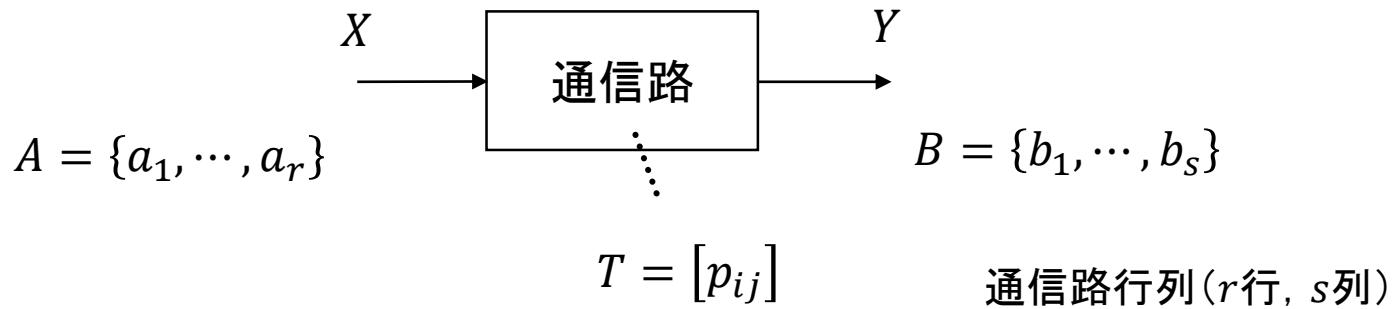
2元対称消失通信路



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

0, 1がXに化ける (=消失する) ことがある。

通信路容量



- (通信路 T による) X と Y の相互情報量:

$$I(X; Y) = \sum_{i=1}^r p_i \left(\sum_{j=1}^s p_{ij} \log_2 \frac{p_{ij}}{q_j} \right) = H(Y) - H(Y|X)$$

- 通信路 T の通信路容量:

$$C = \max_p I(X; Y) \quad \text{ここで, } p = (p_1, \dots, p_r) \text{ は入力の確率分布}$$

まとめ

- 情報量: メッセージが不確かさ(エントロピー)を減らす度合いを定量化
- 不確かなメッセージからでも情報は得られる→相互情報量として定量化
- 相互情報量の性質
- 通信路の働きは相互情報量で説明される.
- 通信路自体は条件付確率を使ってモデル化する→通信路行列
- 基本的な通信路として2元対称通信路と2元対称消失通信路がある.
- 通信路の(理論的)性能は通信路容量で表す.