創想館　14-201

2008-5

Tuesday　SPRING

13:00−14:30

2008年度

「情報工学」

第5回講義

復 習

前回学んだこと

- 覚えていなければいけないこと
- 答えられなければならないこと

✓ エントロピーの‘最大値’を求める－その2

- 微分によって，極値（最大値）を求める.
  - ・ 一般的には無理
  - ・ ３事象のみの場合について求めた

✓ 冗長度 γ

- $\gamma = 1 - (H / H_{max})$
- 通報の‘無駄さ’の度合

✓ 確率過程をマルコフ過程（遷移確率を加味）と考えた場合，エントロピー

- $H_{M1} = -\sum_{i,j=1}^{N} p(i)\, p(j/i) \log_2 p(j/i) \quad (bit)$
- マルコフ過程で考えると，エントロピーは常に小さくなる！

ANIMATION

復 習

前回学んだこと

- 覚えていなければいけないこと
- 答えられなければならないこと

✓ 誤りによる情報量の低減

- $\left[\text{誤りがある場合の情報量}\right] = \left[\text{誤りがない場合の情報量}\right] - \left[\text{誤りによって失われる情報量}\right]$

$\left[\text{誤りによって失われる情報量}\right] = \log_2 \dfrac{1}{\text{その事象が本当に生起　する確率}}$

✓ 誤りがある場合のエントロピー（送信エントロピー）

- $H' = H(x) - H(x/y)$

$$H(x) = -\sum_{i=1}^{N} p(x_i) \log_2 p(x_i)$$

$$H(x/y) = -\sum_{i=1}^{N} p(y_i)\, p(x_j / y_i) \log_2 p(x_j / y_i)$$ あいまい度

ANIMATION

前回学んだこと

復 習

- 覚えていなければいけないこと
- 答えられなければならないこと

ただ，前回の講義は‘失敗’でした！
内容が‘多過ぎ’・・・
多くの学生諸君は，‘消化不良’に陥ってしまったのではないかと思います．　申し訳ありませんでした．

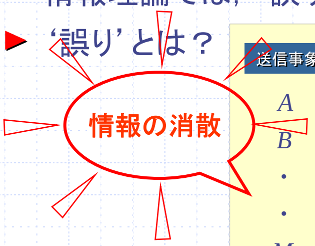
・・・したがって，今日は前回の講義の後半を，
もう一度やり直したいと思います．

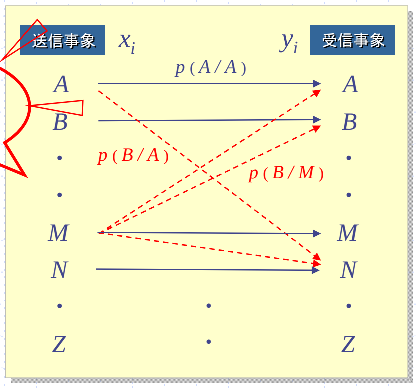
$$H(x) = -\sum_{i=1}^{N} p(x_i)\log_2 p(x_i)$$

$$H(x/y) = -\sum_{i=1}^{N} p(y_i)p(x_j/y_i)\log_2 p(x_j/y_i)$$ あいまい度

§ 3.3　誤りによる情報量の低減

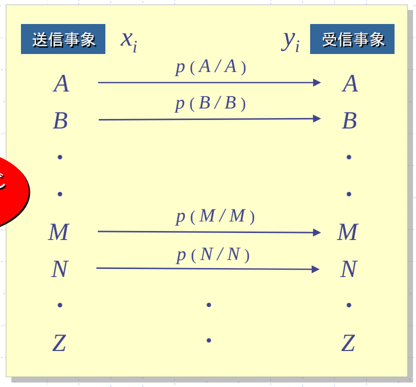
- 発言に，‘誤り’があったり，‘嘘’があったりすると，その発言者の発言は情報量が低下する．　常識！
- 情報理論では，‘誤り’を‘嘘’を区別しない．
- ‘誤り’とは？





§ 3.3　誤りによる情報量の低減

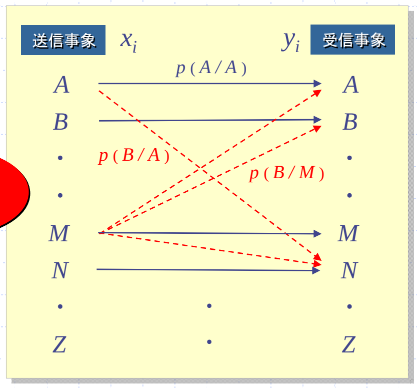
‘誤り’ が生じない場合



[誤りがない場合の送信エントロピー] $= H(x) = -\sum_{i=1}^{N} p(x_i)\log_2 p(x_i)$

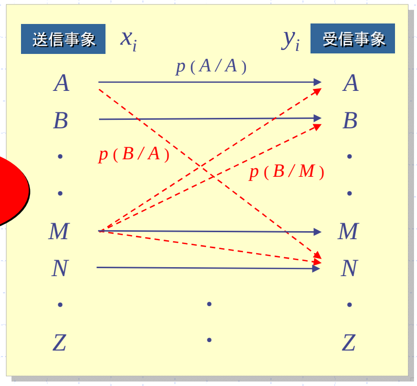
§ 3.3　誤りによる情報量の低減

‘誤り’ が生じる場合



[誤りによって失われる送信エントロピー] $= H(x/y) = -\sum_{i=1}^{N} p(y_i)\,p(x_j/y_i)\log_2 p(x_j/y_i)$

§ 3.3　誤りによる情報量の低減



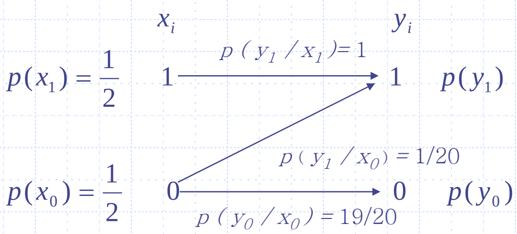
‘誤り’ が生じる場合

［誤りがある場合の送信エントロピー］＝［誤りがない場合の送信エントロピー］－［誤りによって失われる送信エントロピー］

$$\Rightarrow \quad H' = H(x) - H(x/y)$$

Example 3.3.1

各通報の生起確率と送信誤り率（遷移確率）が，下図によって与えられる１，０の通報系列に対し，送信エントロピーを算出しなさい．



例

送信通報 x_i：１０１０１００１０１１１０１００・・・・・

誤　　　　誤

受信通報 y_i：１０１１１００１０１１１０１０１・・・・・

ANIMATION

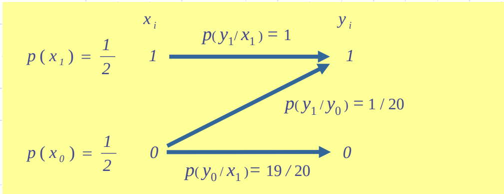
[略解]

$$H(x) = -\sum_{i=1}^{2} p(x_i)\log_2 p(xi) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1\ bit$$

$H(x/y) =$ ？？？　　これを，どう求めるか？

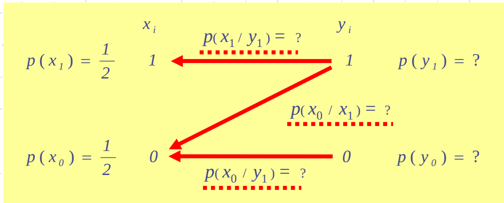
問題から解る情報は･･･

でも，この情報だけから $H(x/y)$ を計算することはできない．



$H(x/y)$ を計算するためには，

$p(x_1/y_1)$, $p(x_0/y_1)$, $p(x_0/y_0)$ を求めなければならない．



[略解]

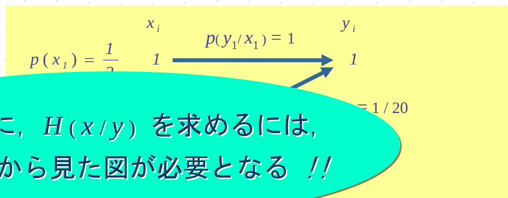
$$H(x) = -\sum_{i=1}^{2} p(x_i)\log_2 p(xi) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1\ bit$$

$H(x/y) =$ ？？？　　これを，どう求めるか？

問題から解る情報は･･･

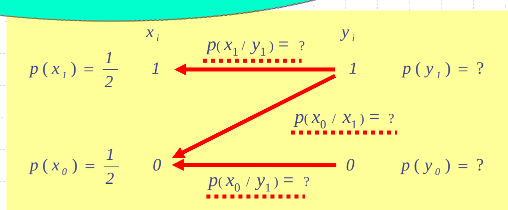
でも，この情報だけから H … を計算すること…

要するに，$H(x/y)$ を求めるには，
受信側から見た図が必要となる !!



$H(x/y)$ を計算するためには，

$p(x_1/y_1)$, $p(x_0/y_1)$, $p(x_0/y_0)$ を求めなければならない．



[略解] ・・・ つづき

まず，$p\,(y_1)$，$p\,(y_0)$ を求めよう．

$$p(y_1) = \underbrace{p(x_1)}_{1/2}\underbrace{p(y_1/x_1)}_{1} + \underbrace{p(x_0)}_{1/2}\underbrace{p(y_1/x_0)}_{5/100} = \frac{105}{200}$$

同様にして，

$$p(y_0) = p(x_1)p(y_0/x_1) + p(x_0)p(y_0/x_0) = \frac{95}{200}$$

次に，受信側から見た各遷移確率は？

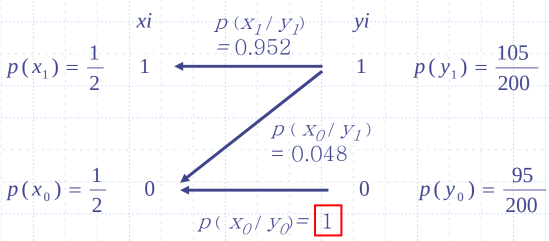
$$\underbrace{p(x_1)}_{1/2}\underbrace{p(y_1/x_1)}_{1} = \underbrace{p(y_1)}_{105/200}\underline{p(x_1/y_1)} \longrightarrow p(x_1/y_1) = 0.952$$

$$\underbrace{p(x_1)}_{1/2}\underbrace{p(y_0/x_1)}_{0} = \underbrace{p(y_0)}_{95/200}\underline{p(x_1/y_0)} \longrightarrow p(x_1/y_0) = 0$$

$$\underbrace{p(x_0)}_{1/2}\underbrace{p(y_1/x_0)}_{5/100} = \underbrace{p(y_1)}_{105/200}\underline{p(x_0/y_1)} \longrightarrow p(x_0/y_1) = 0.048$$

$$\underbrace{p(x_0)}_{1/2}\underbrace{p(y_0/x_0)}_{95/100} = \underbrace{p(y_0)}_{95/200}\underline{p(x_0/y_0)} \longrightarrow p(x_0/y_0) = 1$$

[略解] ・・・ つづき



・・・・・・ 受信側から観察した図

気付いて欲しいこと ・・・

- 送信側で観察した場合には，「１」が信頼できる通報であった．
- しかし，受信側で観察した場合には，「０」が信頼でき通報になっている．

[略解] ・・・ つづき

よって,

$$
\begin{aligned}
H(x/y) = &-\underset{105/200}{p(y_1)}\,\underset{0.952}{p(x_1/y_1)}\log_2 \underset{0.952}{p(x_1/y_1)} \\
&-\underset{105/200}{p(y_1)}\,\underset{0.048}{p(x_0/y_1)}\log_2 \underset{0.048}{p(x_0/y_1)} \\
&-\underset{95/200}{p(y_0)}\,\underset{1}{p(x_0/y_0)}\log_2 \underset{1}{p(x_0/y_0)} \\
&-\underset{95/200}{p(y_0)}\,\underset{0}{p(x_1/y_0)}\log_2 \underset{0}{p(x_1/y_0)} = 0.146\ bit
\end{aligned}
$$

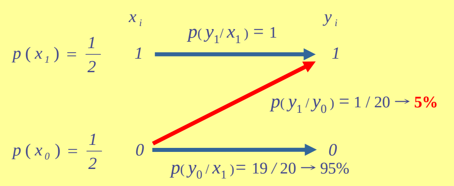
よって,

$$H' = H(x) - H(x/y) = 1 - 0.146 = 0.854\ bit$$

ANIMATION

[略解] ・・・ つづき

よって,



要するに,「0」が5%ほど間違って「1」になってしまうことにより，送信エントロピーは,約15% 減少することになる！

よって,

$$H' = H(x) - H(x/y) = 1 - 0.146 = 0.854\ bit$$

Spring 2008 Prof.M.Nakajima KEIO Univ. Yagami

受信エントロピー

ここまで，通信路において‘誤り’が生じる場合の‘送信情報量’（送信エントロピー）を求めてきた．

受信エントロピーという概念もあって良いはず！

▶ ある！･･･形はこうなる ⇨ $H'' = H(y) - \underline{H(y/x)}$

覚えておくこと！

こちらは、「散布度」という．

ところで．．．

$H(x) \neq H(y)$， $H(x/y) \neq H(y/x)$ であるが，

必ず，

$\underbrace{H(x) - H(x/y)}_{\text{送信エントロピー}} = \underbrace{H(y) - H(y/x)}_{\text{受信エントロピー}}$ となる！

すなわち、
通信路を通して伝送される符号の１符号当たりの平均情報量は、
送信側から見ても、受信側から見ても同じということになる．

ANIMATION

Example 3.3.2：○×問題の採点基準について

問題数 100 問の ○ × 問題があり，全て正答時の ○ × の割合が 50 問 −50 問となることが予め知らされている．P 問正解した受験者には，何点を与えるのが適当か？

情報理論的立場から考えよ！

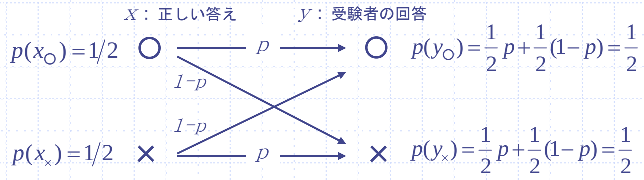
（考え方）

「正答率 p の受験者が，正解１問当たりに送信する平均情報量が，
何ビットになるか？」を考えればよい．

ANIMATION

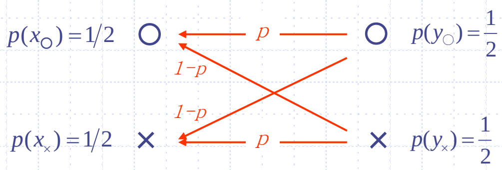
[略解]

まず、正答率 $p= \mathrm{P}/100$ とする. （ P：正答数 ）

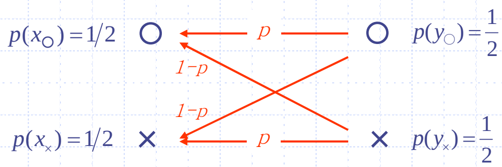


$$H(x)=-\sum_{i=1}^{2}p(x_i)\log_2 p(x_i)=(-\frac{1}{2}\log_2\frac{1}{2})\times 2=1\ bit$$

$H(x/y)=-\sum_{i,j=1}^{2}p(y_i)p(x_j/y_i)\log_2 p(x_j/y_i)$ を求めるためには・・・



[略解] ・・・ つづき



$$H(x)=-\sum_{i=1}^{2}p(x_i)\log_2 p(x_i)=(-\frac{1}{2}\log_2\frac{1}{2})\times 2=1\ bit$$

$$H(x/y)=-\sum_{i,j=1}^{2}p(y_i)p(x_j/y_i)\log_2 p(x_j/y_i)$$

$$=-\underbrace{p(y_{\bigcirc})}_{1/2}\underbrace{p(x_{\bigcirc}/y_{\bigcirc})}_{p}\log_2 \underbrace{p(x_{\bigcirc}/y_{\bigcirc})}_{p}-\underbrace{p(y_{\bigcirc})}_{1/2}\underbrace{p(x_{\times}/y_{\bigcirc})}_{1\text{-}p}\log_2 \underbrace{p(x_{\times}/y_{\bigcirc})}_{1\text{-}p}$$

$$-\underbrace{p(y_{\times})}_{1/2}\underbrace{p(x_{\times}/y_{\times})}_{p}\log_2 \underbrace{p(x_{\times}/y_{\times})}_{p}-\underbrace{p(y_{\times})}_{1/2}\underbrace{p(x_{\bigcirc}/y_{\times})}_{1\text{-}p}\log_2 \underbrace{p(x_{\bigcirc}/y_{\times})}_{1\text{-}p}$$

式を整理すると

$$=-p\log_2 p-(1-p)\log_2(1-p)$$

［ 略解 ］ ・・・ つづき

$$H(x/y) = -p\log_2 p - (1-p)\log_2(1-p)$$

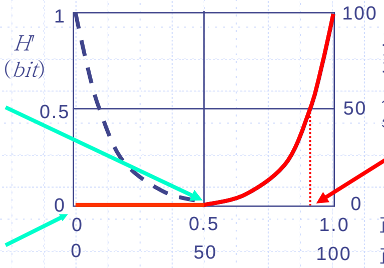
$$\therefore \quad H' = H(x) - H(x/y) = 1 + p\log_2 p + (1-p)\log_2(1-p) \quad bit$$

・・・ これをもって、正答数 P 問に対する、
「送信平均情報量」とすれば良い！

この式をグラフにしてみよう・・・

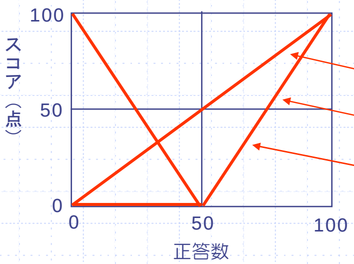
50 問正解というのは、
何も解っていないこと
を意味する！

また、全部間違えるた
めには、全部解ってい
ないと、無理！！



正答率：p
正答数：P （ p= P /100 ）

まとめ



要するに、○×問題の採点は．．．

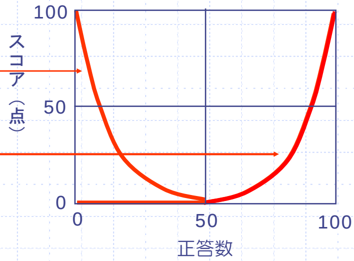
これでは、だめ！

これでも、だめで・・・

これでも、だめ！

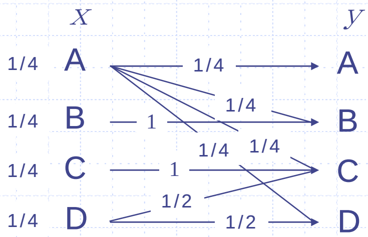
情報理論的には、
これでも良いが ．．．

実用的には、
これが良い！



Example 3.3.3

4つの通報 A,B,C,D の送信に関し、下図のような生起確率と送信誤り率（遷移確率）が与えられている．この通信路において伝送されるエントロピーを算出せよ．



宿題：必ずやってくること！
次回、誰かに回答してもらう！

定期試験は勿論のこと、毎回の授業にも関数電卓を持参のこと！

ANIMATION

2008年度

「情報工学」

第5回講義　おわり