第5章 情報量とひずみ

5.1 情報量の定義

5.1.1 平均符号長の下限としての情報量

情報源を2元系列に変換（符号化）して，その長さで情報の量を計ろうというのである．
長さが最小になるように，2元系列に変換して，その長さを計るべきなのである．

情報量：情報源系列を2元符号に符号化するときの1情報源記号当たりの平均符号長の下限（1情報源記号当たり）単位はbit　Sのエントロピー $H(S)$ で与えられた．

$H(S) = -\sum_{i=1}^{M} p_i \log_2 p_i$（ビット）　この式の対数を自然対数として情報量を計ることがある．

$\log_2 p_i = \frac{\ln p_i}{\ln 2}$．　1ナットは $\frac{1}{\ln 2} = 1.443$ bit　10を底とする対数を用いることがある．
1ハートレーは $\frac{1}{\log_{10} 2} = 3.322$ bit（ハートレー＝ディット＝デシット）

5.1.2 直観的立場からの情報量

(a) 一つの結果を知ったときに得る情報量
確率pの結果の生起を知ったときに，得る情報量を $I(p)$ で表すことにしよう．
(I) $I(p)$ はpの単調減少関数である．(II) AとBは独立，$I(p_A p_B) = I(p_A) + I(p_B)$
(III) $I(p)$ はpの連続関数である．

$I(p) = -a \log_2 p$（a：定数）　$I(\frac{1}{2}) = 1$ のとき，$a=1$，$I(p) = -\log_2 p$

1ビット：確率 $\frac{1}{2}$ で生じる結果　1ナットは確率 $\frac{1}{e}$ で生じる結果　1ハートレーは $\frac{1}{10}$

(b) 平均情報量
記号 $a_1, a_2, \cdots a_M$ 確率 $p_1, p_2, \cdots p_M$ 平均：$\bar{I} = -\sum_{i=1}^{M} p_i \log_2 p_i$ 平均情報量

5.2 エントロピーと情報量．（情報量 $\xleftarrow{定义}$ エントロピー）

エントロピー＝1情報源記号当たりの平均符号長の下限＝情報量．曖昧さを表す尺度
曖昧さ　$H(S) \longrightarrow 0$　$H(S)-0$ は情報量．情報量＝情報を受け取ることによるエントロピーの変化
出力を知っていない　出力を知る

5.2.2 エントロピーの最小値と最大値

$P_i = \frac{1}{M}$ のとき，すなわちエントロピーは，情報源アルファベットの各記号が等確率で発生するとき最大となり，$\log_2 M$ となるのである．　$0 \le H(S) \le \log_2 M$　記憶のある定常M元情報源についても成立する．

$h(S) = \frac{H(S)}{\log_2 M}$ 相対エントロピー　$0 \le h(S) \le 1$

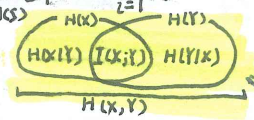
$\rho(S) = 1 - h(S)$ この情報源Sの冗長度（redundancy）と呼ぶ

5.2.3 代表的な系列の数とエントロピー

記憶のない情報源Sの長さn．含まれる記号 a_i の個数 n_i．nが十分大きければ，n_i/n は p_i にきわめて近い値をとる．　$\because n_i \approx np_i$

長さnの代表的系列をσとし，発生確率 $P(\sigma) = \prod_{i=1}^{M} p_i^{n_i}$　$\therefore P(\sigma) \approx \prod_{i=1}^{M} p_i^{np_i} = \prod_{i=1}^{M} 2^{np_i \log_2 p_i}$
代表的系列の数は $1/P(\sigma) = 2^{nH(S)}$　$= 2^{n\sum_{i=1}^{M} p_i \log_2 p_i} = 2^{-nH(S)}$

5.3 相互情報量



相互情報量とエントロピーの関係．

5.3.1 相互情報量の定義

Yで条件をつけたXの条件付エントロピーと呼ばれ．Yを知ったときXについてなお残っている平均の曖昧さの尺度を与える．$H(X|Y)$　　XとYの相互情報量

$H(X|Y) = -\sum_y P(y) \sum_x P(x|y) \log_2 P(x|y)$　　$I(X;Y) = H(X) - H(X|Y)$

5.3.2 相互情報量的性質

$I(X;Y) = \sum_x \sum_y P(x,y) \log_2 \frac{P(x|y)}{P(x)} = \sum_x \sum_y P(x,y) \log_2 \frac{P(x,y)}{P(x)P(y)} = I(Y;X) = H(Y) - H(Y|X)$

$I(X;Y) = H(X) + H(Y) - H(X,Y)$　　$H(X,Y) = -\sum_x \sum_y P(x,y) \log_2 P(x,y)$
$= H(X) - H(X|Y) = H(X) + H(Y) - H(X,Y)$

$I(X;Y) = \lim_{n \to \infty} \frac{1}{n} I(X_n; Y_n)$

5.4 ひずみが許される場合の情報源符号化

5.4.1 情報源符号化におけるひずみ

出力：x　xとyの相違を評価する関数 $d(x,y)$ ひずみ測度　　平均ひずみ $\bar{d}$：$\bar{d} = \sum_x \sum_y d(x,y) P(x,y)$
復号結果：y　$d(x,y) \ge 0$，$x=y$ のとき $d(x,y)=0$　　$P(x,y)$：xとyの結合確率

$d(x,y) = \begin{cases} 0; & x=y \\ 1; & x \ne y \end{cases}$　$\bar{d} = P(1,0) + P(0,1)$　ビット誤り率

$d(x,y) = |x-y|^2$　2乗平均誤差（mean square error）

5.4.2 ひずみが許される場合の情報源符号化定理

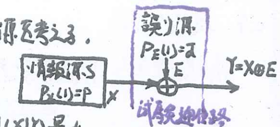
$I(X;Y)$ が同じでも，$\bar{d}$ は同じであるとは限らない．　ある値Dに対し，$\bar{d} \le D$

$R(D) = \min\{I(X;Y)\}$　送らねばならない情報量の最小値，速度・ひずみ関数

平均ひずみ $\bar{d}$ をD以下に抑えるという条件の下で；

$R(D) \le L < R(D) + \varepsilon$（1情報源記号当たりの平均符号長L）

例5.4　1,0を確率p, 1-pで発生する記憶のない2元情報源を考える．



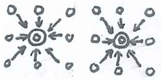
$d(x,y) = \begin{cases} 0; & x=y \\ 1; & x \ne y \end{cases}$　平均ひずみ $\bar{d}$ はビット誤り率

$H(X) = \mathcal{H}(p)$　$I(X;Y) = H(X) - H(X|Y)$．$I(X;Y)$ 最小 → $H(X|Y)$ 最大

$Y = X \oplus E \Rightarrow X = Y \oplus E \Rightarrow H(X|Y) = H(Y \oplus E|Y) = H(E|Y)$（Yの値を知ったときEの曖昧さである）

$H(E|Y) \le H(E) \le \mathcal{H}(\bar{d})$
$H(X|Y) \le \mathcal{H}(\bar{d}) \le \mathcal{H}(D)$　$\Rightarrow I(X;Y) \ge \mathcal{H}(p) - \mathcal{H}(D) = R(D)$

$\therefore R(D) = \mathcal{H}(p) - \mathcal{H}(D)$　送らねばならない情報量の最小値．

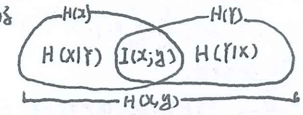


各点は長さnの系列
◎は符号語を表す

5.4.4 ひずみが許される場合の情報源符号化法

情報源S $\xrightarrow{x}$ C_0 への符号化 $x \to w$ $d_n(x,w)$ 最小 $\xrightarrow{w}$ ひずみのない場合の情報源符号化　　$d_n(x,w) = \sum_{i=1}^{n} d(x_i, w_i)$

①平均ひずみができるだけ小さくなるように作るべきである
②1情報源記号当たりの平均符号長を短くできる



第五章总结

5.1 情報量

エントロピー＝1情報源記号当たりの平均符号長の下限　曖昧さを表す尺度
情報量＝情報を受け取ることによるエントロピーの変化　$I(p)$ と書く

相対エントロピー：$h(S) = \frac{H(S)}{\log_2 M}$
冗長度：$\rho(S) = 1 - h(S)$

5.2 相対情報量

Yで条件をつけたXの条件付エントロピーと呼ばれ．Yを知ったときXについてなお残っている平均の曖昧さの尺度を与える　$I(X;Y) = \sum_x \sum_y P(x,y) \log_2 \frac{P(x,y)}{P(x)P(y)} = H(X) - H(X|Y) = H(Y) - H(Y|X)$

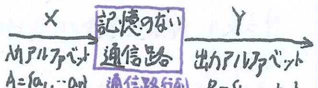
5.3 ひずみ

$d(x,y) = \begin{cases} 0 & x=y \\ 1 & x \ne y \end{cases}$　$\bar{d} = \sum_x \sum_y d(x,y) P(x,y)$　$I(X;Y) = \mathcal{H}(p) - \mathcal{H}(D) = R(D)$
送らねばならない情報量の最小値

第6章 通信路符号化の限界

6.1 通信路容量

入力X 出力Y，相互情報量 $I(X;Y) = \sum_{i=1}^{r} p_i \sum_{j=1}^{s} p_{ij} \log_2 \frac{p_{ij}}{q_j}$



$p_i = P_X(a_i)$　$T = [p_{ij}]$　$p_{ij} = P(b_j|a_i)$　$q_j = P_Y(b_j)$

$q_j = \sum_{i=1}^{r} p_i p_{ij}$　長さn．入力系列 X_n 出力系列 Y_n　P_n を X_n の確率とすれば

通信路容量 $C = \lim_{n \to \infty} \left[\max_{P_n} \left\{ \frac{1}{n} I(X_n; Y_n) \right\} \right]$

6.1.2 記憶のない一様通信路の通信路容量

$C = \log_2 s + \sum_{j=1}^{r} p_{ij} \log_2 p_{ij}$（sは出力系列の長さ）

二元対称通信路

通信路行列：$T = \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$　$C = \log_2 2 + p \log_2 p + (1-p)\log_2(1-p) = 1 - \mathcal{H}(p)$

二元対称消失通信路

通信路行列：$T = \begin{bmatrix} 1-p_x-p & p_x & p \\ p & p_x & 1-p_x-p \end{bmatrix}$

$C = \max_p \{H(Y)\} + (1-p_x-p)\log_2(1-p_x-p) + p_x \log_2 p_x + p \log_2 p$

$C = (1-p_x)\left[1 - \mathcal{H}\left(\frac{p}{1-p_x}\right)\right]$

6.1.3 加法的2元通信路の通信路容量

$I(X;Y) = H(Y) - H(Y|X) = H(Y) - H(X \oplus E|X) = H(Y) - H(E|X)$

$\because$ 誤りEが入力Xと独立である．$I(X;Y) = H(Y) - H(E)$

$P_X(0) = P_X(1) = \frac{1}{2}$ とすれば $H(Y) = 1$．通信路容量 $C = 1 - H(E)$

同じビット誤り率の2元通信路であっても，バースト誤り通信路の方が，ランダム誤り通信路よりも通信路容量は大きく，より多くの情報を伝送できる．

6.2 通信路符号化の基礎概念

6.2.1 通信路符号

Aの元からなる長さnの系列すべての集合を A^n で表す．A^n の中から選ばれた系列の集合C：通信路符号．各系列を符号語と呼ぶ．
通信路符号としては，等長符号が用いられることが多い．
A^n：受信空間　符号語 w_i のまわりの領域 Ω_i は w_i の復号領域と呼ばれ，
受信語がどの復号領域にも属さない領域に入れば，送られた符号語の推定が不可能となる

情報速度 $R = \frac{\log_2 M}{n}$（ビット／記号）　$M = 2^{nR}$　$A = \{a_1, a_2, \cdots a_r\}$

$R_{MAX} = \log_2 r$　$\eta = \frac{R}{R_{max}} \in (0,1)$　Cの冗長度 $\rho = 1 - \eta$

6.2.2 最尤復号法

正しく復号される確率 P_C を最大とする　通信路符号：$C = \{w_1, w_2, \cdots w_M\}$

$P_C(w_i) = \sum_{y \in \Omega_i} P(y|w_i)$　符号語が送られる確率を等しく1/Mであると仮定しよう．

平均の正しく復号される確率　$P_C = \frac{1}{M} \sum_{i=1}^{M} P_C(w_i) = \frac{1}{M} \sum_{i=1}^{M} \sum_{y \in \Omega_i} P(y|w_i)$

最尤復号法は各符号語が等確率で送られるとき，正しく復号される確率 P_C を最大にするという意味で最良の復号法である．

6.3 通信路符号化定理

通信路容量の範囲内で情報を送りながら，いくらでもエラー確率を小さくすることができる

符号作成手順：
1) ビットエラー率pを測定などで求める
2) 必要とする訂正後エラー率εを決める
3) εを実現するための符号間最小ハミング距離を求める
4) 期待する通信路符号の効率を決める
5) それに合う，1符号に含める情報ビット数kと符号長nを決める
6) 符号のk個の情報ビットに対して[illegible]