情報理論 チートシート

---

1. 情報量と情報源

重要公式

- **自己情報量 (Self-Information):** 事象 s_k が生起した時に得られる情報量。
  $I(s_k) = -\log_2 P(s_k)$ [bit]

- **エントロピー (Entropy):** 情報源から得られる平均情報量。
  $H(S) = -\sum_{k=1}^{K} P(s_k) \log_2 P(s_k)$ [bit/symbol]

  - **性質:** $0 \leq H(S) \leq \log_2 K$。$P(s_k)$ が全て等しいとき（等確率のとき）に最大値 $\log_2 K$ をとる。
  - **2元エントロピー関数:** 確率 $p, 1-p$ で生起する2値情報源のエントロピーは $h(p)$ で表される。
    $h(p) = -p \log_2 p - (1-p) \log_2 (1-p)$
    $h(p) = h(1-p)$ であり、$p = 0.5$ のとき最大値 1 をとる。

- **結合エントロピー (Joint Entropy):** $H(X, Y)$
  $H(X, Y) = -\sum_{i=1}^{m} \sum_{j=1}^{n} P(x_i, y_j) \log_2 P(x_i, y_j)$

  - **独立性:** 確率変数 X, Y が独立なとき、$P(x_i, y_j) = P(x_i)P(y_j)$ となるため、$H(X, Y) = H(X) + H(Y)$

- **条件付きエントロピー (Conditional Entropy):** $H(Y|X)$
  $H(Y|X) = -\sum_{i=1}^{m} \sum_{j=1}^{n} P(x_i, y_j) \log_2 P(y_j|x_i)$

- **エントロピーの連鎖律 (Chain Rule for Entropy):**
  $H(X, Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$
  一般形: $H(X_1, ..., X_n) = \sum_{i=1}^{n} H(X_i|X_1, ..., X_{i-1})$

基本用語

- **情報量:** ある事象が起こる確率が低いほど、その事象が起こったときに得られる情報量は大きい。
- **記憶のない情報源:** 各記号の生起が、それ以前に生起した記号とは独立である情報源。

---

2. 情報源符号化

情報源から出力される記号を効率的に表現するための符号化。目標は平均符号長 L をエントロピー $H(S)$ に近づけること。

重要公式

- **平均符号長 (Average Codeword Length):**
  $L = \sum_{k=1}^{K} P(s_k) l_k$
  l_k は記号 s_k の符号語長。

- **クラフトの不等式 (Kraft's Inequality):** 瞬時復号可能な符号が存在するための必要十分条件。
  $\sum_{k=1}^{K} D^{-l_k} \leq 1$
  D は符号アルファベットのサイズ（2元符号なら $D = 2$）。

- **情報源符号化定理**: $H(S) \leq L < H(S) + 1$
  - 平均符号長 L はエントロピー $H(S)$ より小さくはできない。

- **符号化効率 (η)**: $\eta = \frac{H(S)}{L}$

主要アルゴリズム

- **ハフマン符号化 (Huffman Coding):** 最適なプレフィックス符号を構成するアルゴリズム。
  1. 全ての情報源記号を、生起確率の降順に並べる。
  2. 最も確率の低い2つの記号をまとめ、その合計確率を持つ新しい節点（ノード）を作る。
  3. 新しい節点を再び確率のリストに加え、ソートする。
  4. リストに1つの節点しか残らなくなるまで、ステップ2と3を繰り返す。
  5. 完成した木（ハフマン木）の枝に 0 と 1 を割り当て、根から各記号の葉までのパスをたどることで符号語が完成する。

基本用語

- **一意復号可能:** 任意の符号語の系列が、一意に元の記号系列に復号できること。
- **瞬時復号可能 (プレフィックス符号):** どの符号語も、他の符号語の接頭部（プレフィックス）になっていない符号。符号系列の終点を待たずに先頭から一意に復号できる。

---

3. 通信路と相互情報量

重要公式

- **相互情報量 (Mutual Information):** 受信記号 Y を得たことで得られる、送信記号 X に関する情報量。
  - $I(X;Y) = H(X) - H(X|Y)$
  - $I(X;Y) = H(Y) - H(Y|X)$
  - $I(X;Y) = H(X) + H(Y) - H(X,Y)$
  - **性質:**
    - **非負性:** $I(X;Y) \geq 0$。等号成立は X と Y が独立のとき。
    - **対称性:** $I(X;Y) = I(Y;X)$
    - **上限:** $I(X;Y) \leq \min(H(X), H(Y))$

- **通信路容量 (Channel Capacity):** 通信路が誤りなしに伝送できる情報量の最大値。入力確率分布 $P(X)$ を最適化することで求める。
  $C = \max_{P(X)} I(X;Y)$

- **各種通信路の容量**
  - **2元対称通信路 (BSC):** 誤り率 p の通信路。
    $C = 1 - h(p)$

- **2元消失通信路 (BEC):** 消失確率 ϵ の通信路。
$C = 1 - \epsilon$
- **Z通信路:** 0は誤らないが、1が0に誤る確率 p がある通信路。容量は等確率入力では達成されない。$C = \log_2(1 + (1-p)^{1-p}p^p)$

基本用語

- **通信路 (Channel):** 情報を送信側から受信側へ伝達する媒体。
- **通信路行列:** 入力と出力の関係を条件付き確率 $P(y_j|x_i)$ で表した行列。
- **イェンゼンの不等式 (Jensen's Inequality):** 凸関数 f に対して、$E[f(X)] \geq f(E[X])$ が成り立つ。対数関数 $\log(x)$ は凹関数であるため、$E[\log(X)] \leq \log(E[X])$ となる。これは相互情報量の非負性などの証明に用いられる。

---

4. 通信路符号化

通信路で発生する誤りを検出・訂正するための符号化。

重要公式

- **ハミング距離 (Hamming Distance):** 2つの同じ長さの符号語間で、対応する位置の記号が異なる箇所の数。$d(u, v)$ で表す。
- **最小ハミング距離 (d_{min}):** 符号に含まれる全ての異なる符号語対のハミング距離の最小値。
- **誤り検出・訂正能力**
  - e ビットの誤り検出: $d_{min} \geq e + 1$
  - t ビットの誤り訂正: $d_{min} \geq 2t + 1$
- **線形符号 (Linear Codes):** (n, k) 符号。k ビットの情報に $n-k$ ビットの冗長ビットを加えて n ビットの符号語を生成する。
  - **生成行列 (G):** $k \times n$ 行列。情報ベクトル u から符号語 c を生成する。 $c = uG$
  - **パリティ検査行列 (H):** $(n-k) \times n$ 行列。$GH^T = 0$ を満たす。
  - **シンドローム (S):** 受信語 r から計算されるベクトル。 $S = rH^T$

基本用語

- **符号語 (Codeword):** 情報ビットに冗長ビットを付加して作られた伝送単位。
- **シンドローム:** 誤りのパターンを特定するための手がかりとなるベクトル。誤りがなければゼロベクトルになる。
- **伝送速度 (Rate):** k ビットの情報を n ビットの符号語に符号化するときの効率。
$R = k/n$ [bit/symbol]
- **シンドローム**: 誤りのパターンを特定するための手がかりとなるベクトル。誤りがなければゼロベクトルになる。