第二版 §2.2. 随机过程的信息度量.

$$H(X_1\cdots X_{m+n}) \leq H(X_1\cdots X_m) + H(X_{m+1}\cdots X_{m+n}).$$ 记作: $h(m+n) \leq h(m)+h(n)$

引理: $f(n)$满足上式, 半可加. 则

$$\lim_{n\to\infty} \frac{1}{n} f(n) = \inf_n \frac{1}{n} f(n)$$

平稳过程 ⟶ 平稳信源

遍历过程 ⟶ 遍历信源

$\{X_{t_1}, X_{t_2} \cdots X_{t_k}\}$ 与 $\{X_{t_1+h}, \cdots X_{t_k+h}\}$ 同分布

Define. $H_\infty(X) \overset{def}{=} \lim_{n\to\infty} \frac{1}{n} H(X_1\cdots X_n).$

引理

(平稳信源熵率) $= \inf_n \frac{1}{n} H(X_1\cdots X_n)$

$$H'_\infty(X) \overset{def}{=} \lim_{n\to\infty} H(X_n|X_{n-1},\cdots X_1)$$ △.

且 $H_\infty(X) = H'_\infty(X)$

平稳分布: 一种概率分布, 具有某种不变性质的概率分布.

(无记忆信源熵率) $H_\infty(X) = H(X_1)$

(k-阶平稳马氏信源) $H_\infty(X) = H(X_{k+1}|X_k,\cdots X_1)$. $\xrightarrow{k=1}$ $H_\infty(X) = H(X_2|X_1)$.

Define. (冗余度) $\log\|\mathcal{X}\| - H_\infty(X)$. 描述信源输出符号携带信息有效程度

△

§2.3. 渐近等分性质. (考试未涉及).

△

The. 2.3.1. $-\frac{1}{n}\log p(X^n)$ 概率收敛至 $H(X)$. (proof. page. 35).

即 $P_r\{|-\frac{1}{n}\log p(X^n) - H(X)| < \varepsilon\} > 1-\varepsilon.$

$$W_\varepsilon^{(n)} = \{X^n \mid |-\frac{1}{n}\log p(X^n) - H(X)| < \varepsilon.\}.$$ △

弱典型序列集 (ε-典型序列)

$$△\ (1-\varepsilon)2^{n(H(X)-\varepsilon)} \leq \|W_\varepsilon^{(n)}\| \leq 2^{n(H(X)+\varepsilon)}.$$

$$p(X^n) \approx 2^{-nH(X)}, \quad \|W_\varepsilon^{(n)}\| \sim 2^{nH(X)}.$$

意义: $\frac{\|W_\varepsilon^{(n)}\|}{\|\mathcal{X}\|^n} \to 0$, 即大概率出现的序列只有极小一部分

§2.4. $\begin{cases} X^n \in W_\varepsilon^{(n)} \longrightarrow \text{正常编码} \\ X^n \bar{\in} W_\varepsilon^{(n)} \longrightarrow \text{统一编为1. (ERROR). 误差概率 } P_e < \varepsilon. \end{cases}$

记码率：$\frac{1}{n}\log M$，$M = \|W_\varepsilon^{(n)}\|$.　　码率：?

* 若以更广低码率压制，则会使 $P_e \to 1$. 上述已是极限

(信源编码定理). $X_1 \cdots X_n$ 无记忆；

1). 码率 $R = \frac{1}{n}\log M > H(X)$ 时，可以有如上编码且 $P_e \to 0$

2). ———————— $< H(X)$ 时，任以 R 编码有 $P_e \to 1$ (不可用).

* 编码效率 $\eta = \frac{H(X)}{R}$.

§2.5. The. 2.5.1. $X_1 \cdots X_n \cdots$ 独立同分布：$-\frac{1}{n}\log Pr(X_1, \cdots, X_n)$ 以概率1

收敛到 $H(X_1) = H(P)$.

~~The. 2.5.2. $X_1, X_2, \cdots$ 遍历的平稳序列，于~~

The. 2.5.3. $X_1, X_2, \cdots$ 取自有限字母集 $\mathcal{X}$. 平稳、遍历、马氏. 则

$$\lim_{n\to\infty} -\frac{1}{n}\log Pr\{X_1, X_2, \cdots X_n\} = H(X_2 | X_1).$$

(Proof. page. 41).

以概率1成立.

第二版 §3.1. 由 CHAPTER. 只要 $R > H(X)$ 就存在编译方案 s.t. $P_e \to 0$.

- 等长码(分组码)
- 非等长码
  - 奇异码
  - 非奇异码
    - 唯一可译
      - 非即时
      - 即时
    - 非唯一可译

编码: $f: \mathcal{X}^n \longrightarrow \mathcal{U}^k$: $f(x^n) = u^k \in \mathcal{U}^k$ 称为码字. $l(x^n)$ 表示 x^n 的码字长

译码: $\varphi: \mathcal{U}^k \longrightarrow \mathcal{X}^n$

码率(编码速率): $R = \frac{k}{n} \log_2 D$, $D = \|\mathcal{U}\|$. C 称为码字集

$N = \|\mathcal{X}\|$. 则无错编码 f 码率: $R = \frac{k}{n} \log_2 D \geq \log_2 N$ △, 因为 $D^{k} \geq N^{n}$.

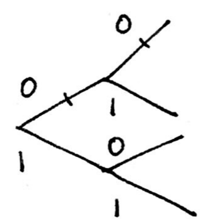
渐近无错 f 码率: $R = \frac{1}{n} \log_2 M \sim H(X)$.

(上述为等长编码)

§3.2. 变长编码.

def. f^* 是 f 的有限扩张码: $f^*(x_1 \cdots x_m) = (f(x_1) \cdots f(x_m))$.

码树构造及即时码:



2进码树

路长为2.

满树 & 非满树

再选

选即时码原则: 选中一个结点, 则其之后所有有均不可

l_{max}: 最大码字长.
l_i: 第 i 层.

有: $\sum_i D^{-l_i} \leq 1$. (Kraft).

(D是进制).

(即: 所有后继结点数 ≥ 既有结点数

$\Longrightarrow$ (Kraft 不等式): 满足 $\sum_i^m D^{-l_i} \leq 1$ 则必存在一个即时码 $(l_1 \cdots l_m)$.
(码长).

反之亦然. 若 $(l_1 \cdots l_m)$ 即时. 则有 $\sum_1^m D^{-l_i} \leq 1$ 成立.

满足Kraft 未必最优； 平均码长 $\sum_i \cancel{p(x_i)}\, p_i l_i = \bar{L}$

为使 $\bar{L}$ 最小：（Larange 乘子法求解）. Page.52.

⇒（最优码长定理）：D进码平均码长 $\bar{L}$： $\bar{L} \ge H_D(X) = -\sum_i^m p_i \log_D p_i$

Proof. page.52. 当且仅当 $D^{-l_i} = p_i$ 时等号成立. $l_i = \lceil \log_D \frac{1}{p_i} \rceil$

按此构造码称为香农码

推广到无穷：$n \to +\infty$ 时最优码平均码长就是 $H_\infty(X)$. （平稳遍历）.

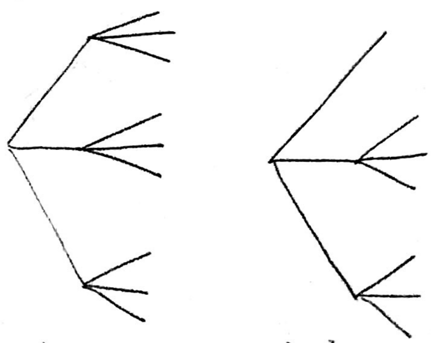
$H(X_1)$ （~~平稳~~独立同分布）

$H(X_2|X_1)$ （平稳马氏）.

def. 冗余度：$\bar{L} - H_D(X)$. or $l^* - H_\infty(X)$.

相对冗余度：$1 - \frac{H_D(X)}{\bar{L}}$ or $1 - \frac{H_\infty(X)}{l^*}$

满树≠完全树（只要求有的结点满分支）



满树 > 完全树

第五版 §4.1 信道编码器编码速率：$R=\frac{1}{n}\log\|W\|$ 码率上限即是信道容量

(f,g). $\begin{cases} f: W \longrightarrow X^n \\ g: Y^n \longrightarrow W \end{cases}$

$P_e = Pr\{g(f(W)) \neq W\}$

$$I(X;Y) = \sum\sum p(x,y)\log\frac{p(x,y)}{p(x)\cdot p(y)} = \sum p(x)\sum \underset{Q}{p(y|x)}\cdot\log\frac{\underset{Q}{p(y|x)}}{\sum_x p(x)\underset{Q}{p(y|x)}}$$

$= I(p;Q)$.

（离散无记忆）

信道容量：$C = \max I(X;Y)$. 有记忆：$\lim\limits_{N\to\infty}\frac{1}{N}\max I(X_1\cdots X_N; Y_1\cdots Y_N)$

对称信道：每一行都是其他行的置换.
每一列都是其他列的置换. $C=\log\|Y\|-h(\underline{r})$

弱对称信道：每一行都是其他行的置换.
每一列元素和相等. $C=\log\|Y\|-h(\underline{r})$

信道容量计算：$p^*(x)$ 是使 $I(p,Q)$ 即 $I(X,Y)$ 最大的充要条件：

$$D(Q(y|x)\,\|\,p(y))\Big|_{p^*(x)} \begin{cases} = C, & p^*(x)>0 \\ \leq C, & p^*(x)=0 \end{cases}$$ （边界条件）

其中 $p(y)=\sum_x p(x)\cdot Q(y|x)$.

* $p^*(x)$ 不一定唯一，但 $p^*(y)$ 是唯一的

信道容量性质：$C\geq 0$

$C\leq\log\|X\|$ $(I(X;Y)\leq H(X)\leq\log\|X\|)$

$C\leq\log\|Y\|$.

§4.3. $[\mathcal{X}, Q(y|x), \mathcal{Y}]$, (M, n):

1. M 消息集 $\mathcal{W} = \{1, 2, \cdots M\}$.
2. 编码 $f: \mathcal{W} \to \mathcal{X}^n$. $\{X^n(1), \cdots X^n(M)\}$
3. 译码 $g: \mathcal{Y}^n \mapsto \mathcal{W}$
4. 码率 $R = \frac{1}{n}\log M$. (码率)

码率出现在 $\begin{cases} 1. 普通的 \frac{1}{n}\log M \\ 2. 分组码的 \frac{K}{n}\log_2 D \end{cases}$

误差: $\lambda_i = Pr\{g(Y^n) \neq i \mid X^n = x^n(i)\}$ (第 i 个信息: $i \in \mathcal{W}$.).

最大误差: $\lambda^{(n)} = \max_{i \in \mathcal{W}} \lambda_i$, 平均: $P_e = \frac{1}{M}\sum_{i}^{M} \lambda_i$

码率 R 可达: $\lim_{n\to\infty} \lambda^{(n)} = 0$. $\underline{(2^{nR}, n)码}$. (M, n). $\underline{R = \frac{1}{n}\log M}$.

~~与容!~~

与信道容量一致 C. $R < C$ 则必可达, R 可达 R 必有 $R \leq C$.

(Ref. 渐近等分性质).

def. n 长联合典型序列 $\{(x^n, y^n)\}$ 集 $W_\varepsilon^{(n)}$ 为

考什么?

$$W_\varepsilon^{(n)} = \left\{ (x^n, y^n) \in \mathcal{X}^n, \mathcal{Y}^n : \begin{array}{l} \left| -\frac{1}{n}\log p(x^n) - H(X) \right| < \varepsilon \\ \left| -\frac{1}{n}\log p(y^n) - H(Y) \right| < \varepsilon \\ \left| -\frac{1}{n}\log p(x^n, y^n) - H(X, Y) \right| < \varepsilon \end{array} \right\}$$