

信息论第三单元复习：信源编码

基本概念

1. 请给出**信源编码 (Source Code)** 的严格定义（包含映射关系和符号说明）。
2. 定义信源编码的**期望长度 (Expected Length)** $L(C)$ ，并写出其数学表达式。
3. 请解释以下编码类型的区别：
 - (a) 奇异码 (Singular Code)
 - (b) 非奇异码 (Non-singular Code)

(c) 唯一可译码 (Uniquely Decodable Code)

(d) 前缀码 (Prefix Code)

Kraft 不等式与 McMillan 定理

4. 请表述 **Kraft 不等式** (离散字母表 D 情况)。

5. 证明：对于任意前缀码，码长 $\{l_i\}$ 必须满足 Kraft 不等式。

6. 证明：若码长 $\{l_i\}$ 满足 Kraft 不等式，则存在对应的前缀码。

7. 请表述推广的 Kraft 不等式（无限字母表 D 情况）。
8. 请表述并证明 McMillan 定理（唯一可译码与 Kraft 不等式的关系）。

最优码长界与香农码

9. 证明最优码长的下界： $L \geq H_D(X)$ （其中 $H_D(X)$ 是以 $|D|$ 为底的信源熵）。

10. 证明等号成立条件：当 $p_i = |D|^{-l_i}$ 时， $L = H_D(X)$ 。

11. 定义香农码 (Shannon Code) 并证明其期望长度满足：

$$H_D(X) \leq L < H_D(X) + 1$$

哈夫曼编码

12. 描述哈夫曼编码 (Huffman Code) 的构造算法。

13. 证明哈夫曼编码满足以下最优性条件：

(a) 若 $p_j > p_k$, 则 $l_j \leq l_k$

(b) 两个最长码字长度相同

(c) 两个最长码字仅在最后一位不同

14. 证明哈夫曼编码是最优前缀码（使用归纳法）。

15. 比较香农码与哈夫曼码的优缺点（期望长度、构造复杂度等）。