第三章 渐进均分性

- 随机变量的收敛：给定一个随机变量序列 $X_1, X_2, \cdots$，序列收敛于随机变量X有如下三种情况
  - 如果对任意的 $\epsilon > 0, \Pr\{|X_n - X| > \epsilon\} \to 0$，则称为依概率收敛
  - 如果 $E(X_n - X)^2 \to 0$，则称为均方收敛
  - 如果 $\Pr\{\lim_{n\to\infty} X_n = X\} = 1$，则称为以概率1（或称几乎处处）收敛。

渐进均分性定理

- 弱大数定理：针对独立同分布（i.i.d.）的随机变量 $X_1, X_2, \cdots$，当n很大时，

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to EX \text{ 依概率}$$

- 渐进均分性定理（Asymptotic Equipartition Property，AEP）

若 $X_1, X_2, \cdots, X_n$ 为i.i.d.~p(x)，则

$$-\frac{1}{n}\log p(X_1, X_2, \cdots, X_n) \to H(X) \text{ 依概率}$$

体现AEP的例子

例

$$X=\begin{cases}1 & 概率为p\\ 0 & 概率为1-p\end{cases}$$

- $X_1,X_2,\cdots,X_n$ 为i.i.d.~p(x)，则 $x_1,x_2,\cdots,x_n$ 出现的概率为 $\prod_{i=1}^{n}p(x_i)$
- 预测出实际观测到的序列的概率：

$$p(X_1,X_2,\cdots,X_n)=p^{\sum X_i}(1-p)^{n-\sum X_i}\rightarrow 2^{-nH(p)}$$

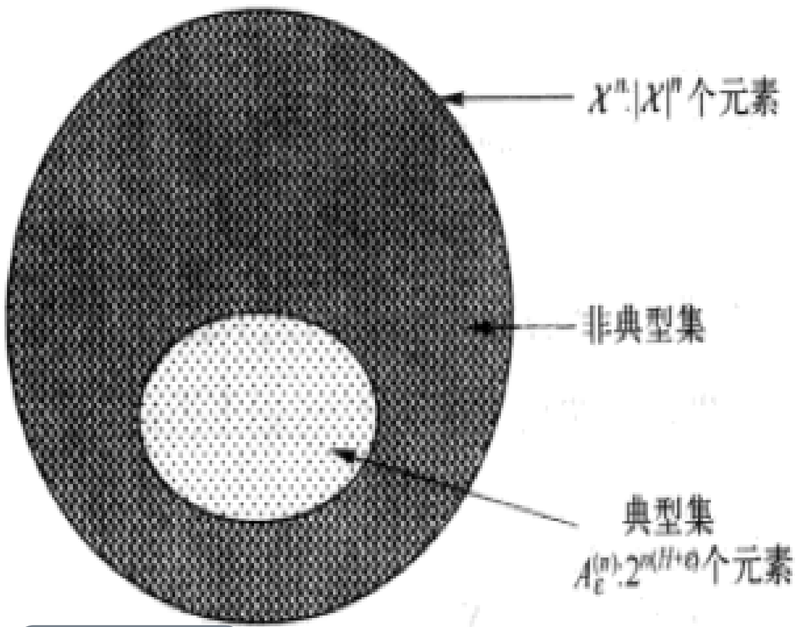
典型集（typical set）

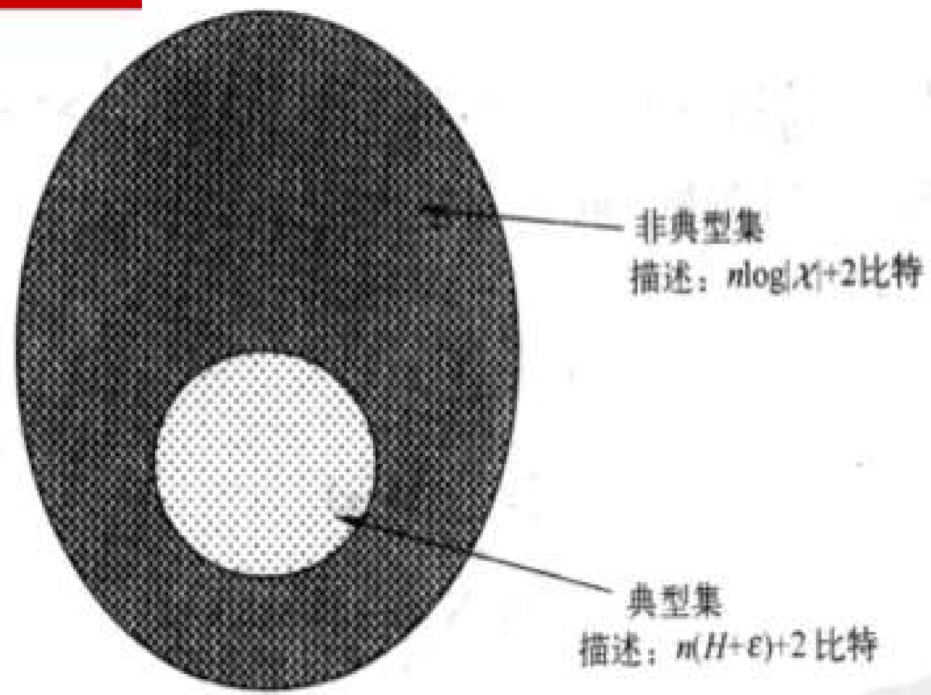
- 关于p(x)的典型集 $A_\epsilon^{(n)}$ 是序列 $x_1, x_2, \cdots, x_n \in \chi^n$ 的集合，且满足性质

$$2^{-n(H(X)+\epsilon)} \le p(x_1, x_2, \cdots, x_n) \le 2^{-n(H(X)-\epsilon)}$$

- 典型集的性质：
  - 如果 $(x_1, x_2, \cdots, x_n) \in A_\epsilon^{(n)}$, $H(X) - \epsilon \le -\frac{1}{n}\log p(x_1, \cdots, x_n) \le H(X) + \epsilon$
  - 当n充分大时，$\Pr\{A_\epsilon^{(n)}\} > 1 - \epsilon$
  - $|A_\epsilon^{(n)}| \le 2^{n(H(X)+\epsilon)}$，其中|A|表示集合中元素个数
  - 当n充分大时，$|A_\epsilon^{(n)}| \ge (1-\epsilon)2^{n(H(X)-\epsilon)}$

AEP的推论：数据压缩





定理 设 X^n 为服从 $p(x)$的i.i.d序列，$\varepsilon>0$，则存在一个编码将长度为 n 的序列x^n映射为比特串，使得映射是1－1的(因而可逆)，且对于充分大的 n，有

$$E\left[\frac{1}{n}l(X^n)\right]\leqslant H(X)+\varepsilon \tag{3-23}$$