课堂小测

考虑以下随机变量分布：

$$\boldsymbol{p}_X = \left[\frac{1}{21}, \frac{1}{21}, \frac{2}{21}, \frac{4}{21}, \frac{6}{21}, \frac{7}{21}\right].$$

(a) 进行二元 Huffman 编码;

(b) 进行三元 Huffman 编码;

(c) 分别计算 (a) 和 (b) 码率。

向量信源编码定理

1. 向量信源编码
2. 大数定律
3. 样本熵和典型集
4. 渐近均分性（AEP）
5. 向量信源编码定理的证明

注：典型集和渐近均分性是信息论的基础

主要内容：

- 描述一个随机向量的**最小无损率**：$H(\boldsymbol{X})$
- 简单、渐近最优"**典型集**"编码
- **渐近均分性**（AEP）证明

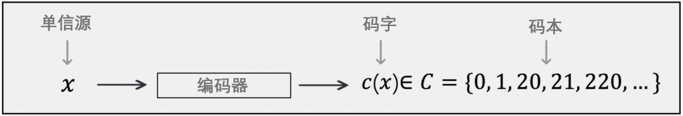
向量信源编码定理

1. 向量信源编码
2. 大数定律
3. 样本熵和典型集
4. 渐近均分性（AEP）
5. 向量信源编码定理的证明

注：典型集和渐近均分性是信息论的基础

向量信源编码

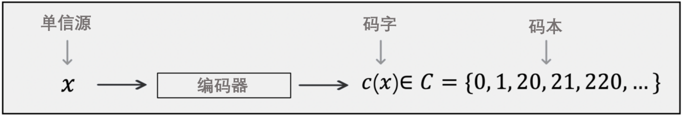
- **单信源编码**



- 输入：单随机变量
- 最优方案：Huffman 编码
- 最优编码 C^*：$H(X) \leq L(C^*) < H(X) + 1$

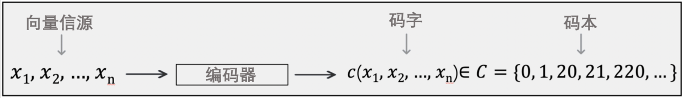
向量信源编码

- 单信源编码



  - 输入：单随机变量
  - 最优方案：Huffman 编码
  - 最优编码 C^*：$H(X) \leq L(C^*) < H(X) + 1$

- 向量信源编码



  - 输入：n 个随机变量组成的序列
  - 最优方案：“典型集”编码（“渐近均分性”）
  - 最优编码 C^*：$H(X) \leq L(C^*) < H(X) + \epsilon, \forall \epsilon > 0$

向量信源编码的符号表示

令 C 为信源编码码本，对一个信源符号 $\boldsymbol{x} \in \mathcal{X}^n$：

- 码字：$C(\boldsymbol{x})$
- 码长：$\ell(\boldsymbol{x})$
- 码率：

$$R(C) = \frac{1}{n} \sum_{\boldsymbol{x} \in \mathcal{X}^n} p_{\boldsymbol{X}}(\boldsymbol{x}) \ell(\boldsymbol{x})$$

单变量情况：$L(C) = \sum_{x \in \mathcal{X}} p_X(x) \ell(x)$

- 单信源编码是向量信源编码 $n = 1$ 时的特例
- 码率 R 越低，描述随机变量序列 $\boldsymbol{x}$ 所需比特数越少

向量信源编码定理

定理（向量信源编码）：令 $\boldsymbol{X} = (X_1, \cdots, X_n)$ 为 n 个独立同分布的随机变量，$X_i \sim p_x$ 的熵为 $H(X)$。则，对任意 $\epsilon > 0$，当 n 足够大时，存在码率为 R 的向量信源编码满足：

$$H(X) \leq R < H(X) + \epsilon.$$

即，当 $n \to \infty$ 时，$R \to H(X)$.

通过单信源编码定理证明

回顾:

定理 (单信源编码定理): 令 L^* 为一个**最优** d 元码本 C^* 的期望码长，则

$$H_d(X) \leq L^* < H_d(X) + 1.$$

等号成立条件为其概率分布是 d **进制**。

通过单信源编码定理证明

回顾:

定理 (单信源编码定理): 令 L^* 为一个**最优**d 元码本 C^* 的期望码长，则

$$H_d(X) \leq L^* < H_d(X) + 1.$$

等号成立条件为其概率分布是 d **进制**。

将 $X = [X_1, X_2, \cdots, X_n]$ **视为一个单随机变量**:

例: $X_i \in \{0, 1\}$，概率为 $\{p_0 = 0.3, p_1 = 0.7\}$，则，

$$\boldsymbol{X} = [X_1, X_2] \in \{00, 01, 10, 11\} \sim \{a, b, c, d\}$$

各元素概率为

$$\{p_a = 0.09, p_b = 0.21, p_c = 0.21, p_d = 0.49\}$$

通过单信源编码定理证明

因此，

$$H_d(\boldsymbol{X}) \leq nL_n^* < H_d(\boldsymbol{X}) + 1,$$

通过单信源编码定理证明

因此，

$$H_d(\boldsymbol{X}) \leq nL_n^* < H_d(\boldsymbol{X}) + 1,$$

$$\frac{1}{n}H_d(\boldsymbol{X}) \leq L_n^* < \frac{1}{n}H_d(\boldsymbol{X}) + \frac{1}{n},$$

通过单信源编码定理证明

因此，

$$H_d(\boldsymbol{X}) \le nL_n^* < H_d(\boldsymbol{X}) + 1,$$

$$\frac{1}{n}H_d(\boldsymbol{X}) \le L_n^* < \frac{1}{n}H_d(\boldsymbol{X}) + \frac{1}{n},$$

$$H(X) \le R < H(X) + \epsilon,$$

其中 $R = L_n^*$, $H(X) = \frac{1}{n}H_d(\boldsymbol{X})$ （$\boldsymbol{X}$ 独立同分布），$\epsilon = 1/n$.

注：可用 Huffman 编码，但复杂度为指数级。

- 单信源编码定理 $\stackrel{\text{证明}}{\Longrightarrow}$ 向量信源编码定理

- 单信源编码定理 $\overset{\text{证明}}{\Longrightarrow}$ 向量信源编码定理
- 但单信源编码定理的证明比较复杂

- 单信源编码定理 $\xRightarrow{\text{证明}}$ 向量信源编码定理
- 但单信源编码定理的证明比较复杂
- 利用**典型集**和**渐近均分性**给出**简易证明**

向量信源编码定理

1. 向量信源编码
2. 大数定律
3. 样本熵和典型集
4. 渐近均分性（AEP）
5. 向量信源编码定理的证明

注：典型集和渐近均分性是信息论的基础

样本均值和数学期望

假设 $X_1, X_2, \cdots, X_n$ 是 n 维独立同分布（IID）随机变量:

- **数学期望 & 方差**，由于 X_i 独立同分布，期望和方差:

$$\mathrm{E}[X] = \sum_{x \in \mathcal{X}} x\, p(x), \quad \mathrm{Var}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$$

样本均值和数学期望

假设 $X_1, X_2, \cdots, X_n$ 是 n 维独立同分布（IID）随机变量:

- **数学期望 & 方差**，由于 X_i 独立同分布，期望和方差:

$$\mathrm{E}[X] = \sum_{x \in \mathcal{X}} x\, p(x), \quad \mathrm{Var}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$$

- **样本均值**$\bar{X}_n$ 为

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

样本均值和数学期望

假设 $X_1, X_2, \cdots, X_n$ 是 n 维独立同分布（IID）随机变量：

- **数学期望 & 方差**，由于 X_i 独立同分布，期望和方差：

$$\mathrm{E}[X] = \sum_{x \in \mathcal{X}} x\, p(x), \quad \mathrm{Var}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$$

- **样本均值**$\bar{X}_n$ 为

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

- 由于X_i **是随机变量**，$\bar{X}_n$ **同样也是随机变量**，且

$$\mathrm{E}[\bar{X}_n] = \mathrm{E}[X], \qquad \mathrm{Var}[\bar{X}_n] = \frac{1}{n}\mathrm{Var}[X] \xrightarrow{n \to \infty} 0$$

样本均值和数学期望

假设 $X_1, X_2, \cdots, X_n$ 是 n 维独立同分布（IID）随机变量:

- **数学期望 & 方差**，由于 X_i 独立同分布，期望和方差:

$$\mathrm{E}[X] = \sum_{x \in \mathcal{X}} x\, p(x), \quad \mathrm{Var}[X] = \mathrm{E}[(X - \mathrm{E}[X])^2] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$$

- **样本均值**$\bar{X}_n$ 为

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

- 由于X_i **是随机变量**，$\bar{X}_n$ **同样也是随机变量**，且

$$\mathrm{E}[\bar{X}_n] = \mathrm{E}[X], \qquad \mathrm{Var}[\bar{X}_n] = \frac{1}{n}\mathrm{Var}[X] \xrightarrow{n \to \infty} 0$$

- 若 X_i 服从伯努利（两点）分布，即 $X_i \sim B(1, p)$，则

$$\Pr(\bar{X}_n = z/n) = \binom{n}{z} p^z (1-p)^{n-z}, \quad z = 0, 1, \cdots, n$$

样本均值实例

- 考虑 $n = 15, p = 1/3$ 的两点分布随机向量 (二元向量):

$$\boldsymbol{X} = X_1, X_2, \cdots, X_{15}.$$

其数学期望是多少? $\Rightarrow 1/3$。

样本均值实例

- 考虑 $n=15, p=1/3$ 的两点分布随机向量（二元向量）：

$$\boldsymbol{X}=X_1, X_2, \cdots, X_{15}.$$

其数学期望是多少？$\Rightarrow 1/3$。

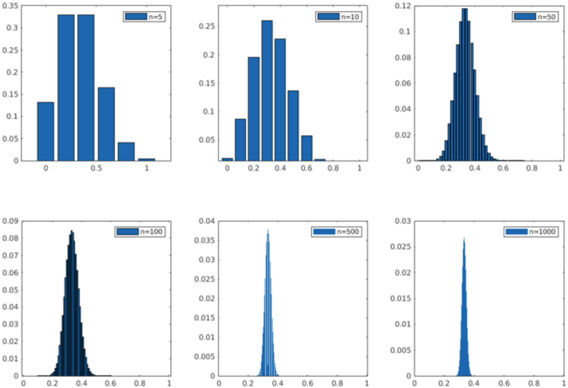
- 样本均值是什么？执行多次随机测试，结果如下：

$$
\begin{array}{l}
0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 1 \rightarrow \text{ave} = 0.266667 \\
1\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0 \rightarrow \text{ave} = 0.4 \\
0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0 \rightarrow \text{ave} = 0.2 \\
1\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1 \rightarrow \text{ave} = 0.466667 \\
0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 1\ 0 \rightarrow \text{ave} = 0.266667 \\
1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0 \rightarrow \text{ave} = 0.2 \\
0\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 1\ 0\ 1\ 0\ 0\ 0\ 0 \rightarrow \text{ave} = 0.266667 \\
1\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0 \rightarrow \text{ave} = 0.333333 \\
0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0 \rightarrow \text{ave} = 0.266667
\end{array}
$$

两点分布样本均值的概率分布

$$\Pr(\bar{X}_n = \frac{z}{n}) = \binom{n}{z} p^z (1-p)^{n-z}$$

$p = 1/3$ 时，期望 $\mathrm{E}[\bar{X}_n] = \mathrm{E}[X] = 1/3$，样本均值的概率分布函数：



$$n \uparrow \Rightarrow \mathrm{Var}[\bar{X}_n] \downarrow \text{ 且 } \lim_{n\to\infty} \mathrm{Var}[\bar{X}_n] \to 0$$

样本均值离数学期望有多近?

- **定义**：考虑随机序列 $\boldsymbol{X}$，若其样本均值 $\bar{X}_n$ 满足

$$|\bar{X}_n - \mathrm{E}[X]| \leq \epsilon,$$

其中 ϵ 为常数，则称其样本均值 $\bar{X}_n$ 与数学期望 $\mathrm{E}[X]$ "ϵ **接近**"。

样本均值离数学期望有多近?

- **定义**：考虑随机序列 $\boldsymbol{X}$，若其样本均值 $\bar{X}_n$ 满足

$$|\bar{X}_n - \mathrm{E}[X]| \leq \epsilon,$$

其中 ϵ 为常数，则称其样本均值 $\bar{X}_n$ 与数学期望 $\mathrm{E}[X]$ “ϵ **接近**”。

- 由于 $\bar{X}_n$ 是随机变量，我们关心：

$$\Pr\left(|\bar{X}_n - \mathrm{E}[X]| \leq \epsilon\right)$$

样本均值离数学期望有多近?

- **定义**：考虑随机序列 $\boldsymbol{X}$，若其样本均值 $\bar{X}_n$ 满足

$$|\bar{X}_n - \mathrm{E}[X]| \leq \epsilon,$$

其中 ϵ 为常数，则称其样本均值 $\bar{X}_n$ 与数学期望 $\mathrm{E}[X]$ "ϵ **接近**"。

- 由于 $\bar{X}_n$ 是随机变量，我们关心：

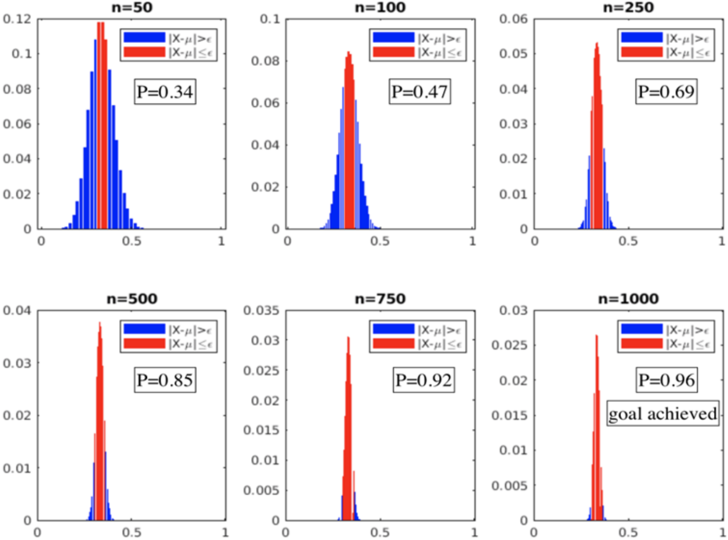
$$\Pr\left(|\bar{X}_n - \mathrm{E}[X]| \leq \epsilon\right)$$

- 直接计算这个概率非常困难，我们考虑其下界 q:

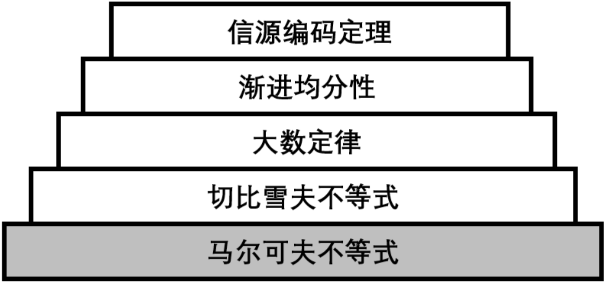
$$\Pr\left(|\bar{X}_n - \mathrm{E}[X]| \leq \epsilon\right) \geq q$$

n 需要多大呢？

令 $p=\frac{1}{3}, \epsilon=0.03$，为使 $\Pr(|\bar{X}_n-\mu|\le\epsilon)>0.95$，$n$ 需要多大呢？



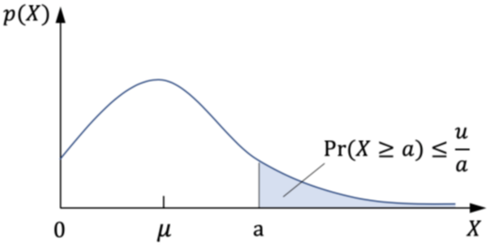
"信息论是大数定律的巧妙应用"



马尔可夫不等式 (Markov Inequality)

马尔可夫不等式：X 是一个**非负**随机变量且 $a > 0$，则

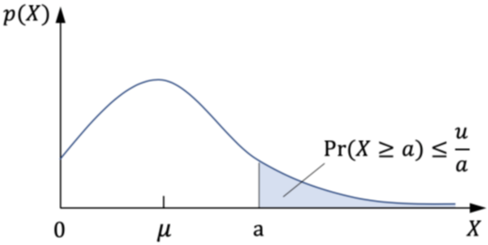
$$\Pr(X \geq a) \leq \frac{\mu}{a}, \quad \mu = \mathrm{E}[X]$$



马尔可夫不等式 (Markov Inequality)

马尔可夫不等式：X 是一个**非负**随机变量且 $a > 0$，则

$$\Pr(X \geq a) \leq \frac{\mu}{a}, \quad \mu = \mathrm{E}[X]$$

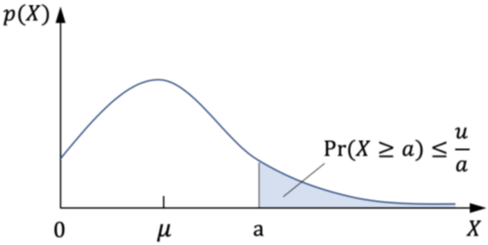


简易证明：$\mu = \mathrm{E}[X] = \sum_x p(x)x$

马尔可夫不等式 (Markov Inequality)

马尔可夫不等式：X 是一个**非负**随机变量且 $a > 0$，则

$$\Pr(X \geq a) \leq \frac{\mu}{a}, \quad \mu = \mathrm{E}[X]$$

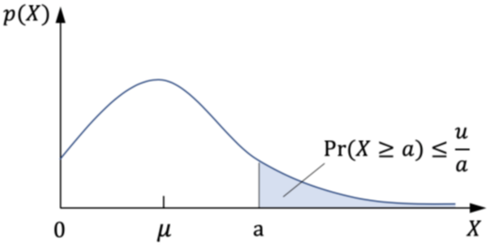


简易证明：$\mu = \mathrm{E}[X] = \sum_x p(x)x \geq \sum_{x \geq a} p(x)x$

马尔可夫不等式 (Markov Inequality)

马尔可夫不等式：X 是一个**非负**随机变量且 $a > 0$，则

$$\Pr(X \geq a) \leq \frac{\mu}{a}, \quad \mu = \mathrm{E}[X]$$

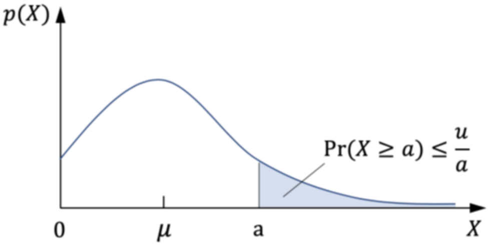


简易证明：$\mu = \mathrm{E}[X] = \sum_x p(x)x \geq \sum_{x \geq a} p(x)x \geq \sum_{x \geq a} p(x)a$

马尔可夫不等式 (Markov Inequality)

马尔可夫不等式：X 是一个**非负**随机变量且 $a > 0$，则

$$\Pr(X \geq a) \leq \frac{\mu}{a}, \quad \mu = \mathrm{E}[X]$$



简易证明：$\mu = \mathrm{E}[X] = \sum_x p(x)x \geq \sum_{x \geq a} p(x)x \geq \sum_{x \geq a} p(x)a = a\Pr(X \geq a)$

“信息论是大数定律的巧妙应用”

切比雪夫不等式 (Chebyshev Inequality)

切比雪夫不等式：考虑一个期望为 μ，方差为 σ^2 的随机变量 X，对任意 $\epsilon > 0$，则

$$\Pr\left(|X-\mu| \geq \epsilon\right) \leq \frac{\sigma^2}{\epsilon^2}.$$

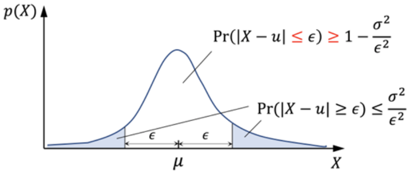
切比雪夫不等式 (Chebyshev Inequality)

切比雪夫不等式：考虑一个期望为 μ，方差为 σ^2 的随机变量 X，对任意 $\epsilon > 0$，则

$$\Pr\left(|X-\mu| \geq \epsilon\right) \leq \frac{\sigma^2}{\epsilon^2}.$$

证明：将$(X-\mu)^2$ **视为期望为** σ^2 **的非负随机变量**，由**马尔可夫不等式**可得

$$\Pr\left((X-\mu)^2 \geq \epsilon^2\right) \leq \frac{\sigma^2}{\epsilon^2}.$$



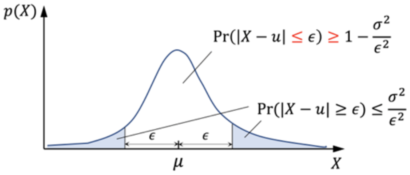
切比雪夫不等式 (Chebyshev Inequality)

切比雪夫不等式：考虑一个期望为 μ，方差为 σ^2 的随机变量 X，对任意 $\epsilon > 0$，则

$$\Pr\left(|X-\mu| \geq \epsilon\right) \leq \frac{\sigma^2}{\epsilon^2}.$$

证明：将$(X-\mu)^2$ **视为期望为** σ^2 **的非负随机变量**，由**马尔可夫不等式**可得

$$\Pr\left((X-\mu)^2 \geq \epsilon^2\right) \leq \frac{\sigma^2}{\epsilon^2}.$$



- 也可以写成如下形式：

$$\Pr\left(|X-\mu| \leq \epsilon\right) \geq 1-\frac{\sigma^2}{\epsilon^2},$$

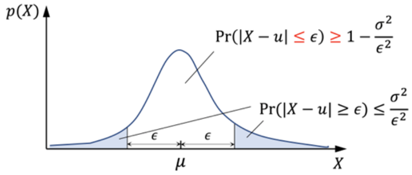
切比雪夫不等式 (Chebyshev Inequality)

切比雪夫不等式：考虑一个期望为 μ，方差为 σ^2 的随机变量 X，对任意 $\epsilon > 0$，则

$$\Pr\left(|X-\mu| \geq \epsilon\right) \leq \frac{\sigma^2}{\epsilon^2}.$$

证明：将$(X-\mu)^2$ **视为期望为** σ^2 **的非负随机变量**，由**马尔可夫不等式**可得

$$\Pr\left((X-\mu)^2 \geq \epsilon^2\right) \leq \frac{\sigma^2}{\epsilon^2}.$$

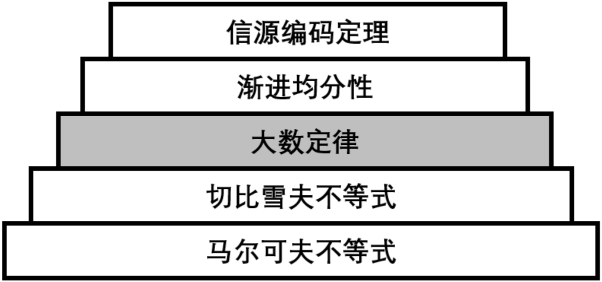


- 也可以写成如下形式：

$$\Pr\left(|X-\mu| \leq \epsilon\right) \geq 1-\frac{\sigma^2}{\epsilon^2},$$

当 $\frac{\sigma^2}{\epsilon^2} \to 0$ 时，$\Pr(|X-\mu| \leq \epsilon) \to 1$.

“信息论是大数定律的巧妙应用”



大数定律 (Law of Large Numbers)

$X_1, X_2, \cdots, X_n$ 为 n 个独立同分布随机变量，且 $\mathrm{Var}[X_i] = \sigma^2$.

弱大数定律：样本均值依概率收敛于数学期望，即对 $\forall \epsilon > 0$,

$$\lim_{n \to \infty} \Pr(|\bar{X}_n - \mathrm{E}[X]| \leq \epsilon) = 1.$$

大数定律 (Law of Large Numbers)

$X_1, X_2, \cdots, X_n$ 为 n 个独立同分布随机变量，且 $\mathrm{Var}[X_i] = \sigma^2$.

弱大数定律：样本均值依概率收敛于数学期望，即对 $\forall \epsilon > 0$,

$$\lim_{n \to \infty} \Pr(|\bar{X}_n - \mathrm{E}[X]| \leq \epsilon) = 1.$$

证明：$\mathrm{E}[\bar{X}_n] = \mathrm{E}[X]$, $\mathrm{Var}(\bar{X}_n) = \frac{\sigma^2}{n} \xrightarrow{n \to \infty} 0$，由切比雪夫不等式，

大数定律 (Law of Large Numbers)

$X_1, X_2, \cdots, X_n$ 为 n 个独立同分布随机变量，且 $\mathrm{Var}[X_i] = \sigma^2$.

弱大数定律：样本均值依概率收敛于数学期望，即对 $\forall \epsilon > 0$,

$$\lim_{n\to\infty} \Pr(|\bar{X}_n - \mathrm{E}[X]| \leq \epsilon) = 1.$$

证明：$\mathrm{E}[\bar{X}_n] = \mathrm{E}[X]$, $\mathrm{Var}(\bar{X}_n) = \frac{\sigma^2}{n} \xrightarrow{n\to\infty} 0$，由切比雪夫不等式，

$$\Pr(|\bar{X}_n - \mathrm{E}[X]| \leq \epsilon) \geq 1 - \frac{\mathrm{Var}(\bar{X}_n)}{\epsilon^2}$$

大数定律 (Law of Large Numbers)

$X_1, X_2, \cdots, X_n$ 为 n 个独立同分布随机变量，且 $\mathrm{Var}[X_i] = \sigma^2$.

> **弱大数定律**：样本均值依概率收敛于数学期望，即对 $\forall \epsilon > 0$,
>
> $$\lim_{n\to\infty} \Pr(|\bar{X}_n - \mathrm{E}[X]| \leq \epsilon) = 1.$$

证明：$\mathrm{E}[\bar{X}_n] = \mathrm{E}[X]$，$\mathrm{Var}(\bar{X}_n) = \frac{\sigma^2}{n} \xrightarrow{n\to\infty} 0$，由切比雪夫不等式，

$$\Pr(|\bar{X}_n - \mathrm{E}[X]| \leq \epsilon) \geq 1 - \frac{\mathrm{Var}(\bar{X}_n)}{\epsilon^2} = 1 - \frac{\sigma^2}{n\epsilon^2}$$

大数定律 (Law of Large Numbers)

$X_1, X_2, \cdots, X_n$ 为 n 个独立同分布随机变量，且 $\mathrm{Var}[X_i] = \sigma^2$.

弱大数定律：样本均值依概率收敛于数学期望，即对 $\forall \epsilon > 0$,

$$\lim_{n\to\infty} \Pr(|\bar{X}_n - \mathrm{E}[X]| \leq \epsilon) = 1.$$

证明：$\mathrm{E}[\bar{X}_n] = \mathrm{E}[X]$, $\mathrm{Var}(\bar{X}_n) = \dfrac{\sigma^2}{n} \xrightarrow{n\to\infty} 0$，由切比雪夫不等式，

$$\Pr(|\bar{X}_n - \mathrm{E}[X]| \leq \epsilon) \geq 1 - \frac{\mathrm{Var}(\bar{X}_n)}{\epsilon^2} = 1 - \frac{\sigma^2}{n\epsilon^2} \xrightarrow{n\to\infty} 1.$$

向量信源编码定理

1. 向量信源编码
2. 大数定律
3. **样本熵和典型集**
4. 渐近均分性（AEP）
5. 向量信源编码定理的证明

注：典型集和渐近均分性是信息论的基础

样本熵 (Sample Entropy)

定义：联合概率分布函数为 $P_{\boldsymbol{X}}$ 的随机序列 $\boldsymbol{X}$，其**样本熵**为

$$\bar{H}(\boldsymbol{X}) = -\frac{1}{n}\log p_{\boldsymbol{X}}$$

样本熵 (Sample Entropy)

定义：联合概率分布函数为 $P_{\boldsymbol{X}}$ 的随机序列 $\boldsymbol{X}$，其**样本熵**为

$$\bar{H}(\boldsymbol{X}) = -\frac{1}{n}\log p_{\boldsymbol{X}}$$

当 $X_1, X_2, \cdots, X_n$ 独立同分布，

$$\bar{H}(\boldsymbol{X}) = -\frac{1}{n}\sum_{i=1}^{n}\log p_{X_i}$$

样本熵 (Sample Entropy)

定义：联合概率分布函数为 $P_{\boldsymbol{X}}$ 的随机序列 $\boldsymbol{X}$，其**样本熵**为

$$\bar{H}(\boldsymbol{X}) = -\frac{1}{n}\log p_{\boldsymbol{X}}$$

当 $X_1, X_2, \cdots, X_n$ 独立同分布，

$$\bar{H}(\boldsymbol{X}) = -\frac{1}{n}\sum_{i=1}^{n}\log p_{X_i}$$

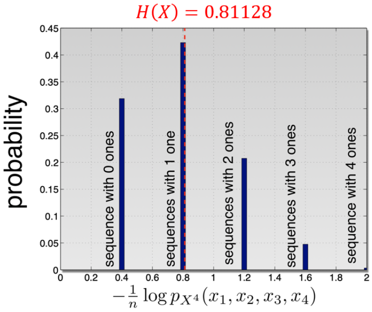
注意：与样本均值和数学期望类似：

- **样本熵**是一个随机变量
- **真实熵**是一个常数

样本熵实例

- 考虑独立同分布的二元随机序列 $\boldsymbol{X} = X_1, X_2, X_3, X_4$，且 $p_0 = \frac{3}{4}, p_1 = \frac{1}{4}$。
- 样本熵的分布（样本总数为 $2^4 = 16$）

x_1, x_2, x_3, x_4	$p_{\mathbf{X}}(\mathbf{x})$	sample entropy $-\frac{1}{n}\log p_{\mathbf{X}}(\mathbf{x})$
0 0 0 0	0.31641	0.41504
1 0 0 0	0.10547	0.81128
0 1 0 0	0.10547	0.81128
0 0 1 0	0.10547	0.81128
0 0 0 1	0.10547	0.81128
1 1 0 0	0.03516	1.20752
1 0 1 0	0.03516	1.20752
0 1 1 0	0.03516	1.20752
1 0 0 1	0.03516	1.20752
0 1 0 1	0.03516	1.20752
0 0 1 1	0.03516	1.20752



典型集 (Typical Set)

定义：令 ϵ 为任意正数，则典型集 $\mathcal{T}_\epsilon^{(n)}$ 是所有**样本熵与真实熵**"ϵ **接近**"的离散无记忆随机序列 $\boldsymbol{X} \in \mathcal{X}^n$ 的集合，即

$$\mathcal{T}_\epsilon^{(n)} = \{\boldsymbol{X} \in \mathcal{X}^n : |\bar{H}(\boldsymbol{X}) - H(X)| < \epsilon\}$$

典型集 (Typical Set)

定义：令 ϵ 为任意正数，则典型集 $\mathcal{T}_\epsilon^{(n)}$ 是所有**样本熵与真实熵**"ϵ **接近**"的离散无记忆随机序列 $\boldsymbol{X} \in \mathcal{X}^n$ 的集合，即

$$\mathcal{T}_\epsilon^{(n)} = \{\boldsymbol{X} \in \mathcal{X}^n : |\bar{H}(\boldsymbol{X}) - H(X)| < \epsilon\}$$

- 典型集的元素个数表示为 $|\mathcal{T}_\epsilon^{(n)}|$;
- 信源集合 $\mathcal{X}^n$ 可以分为典型集 $\mathcal{T}_\epsilon^{(n)}$ 及其补集 $\bar{\mathcal{T}}_\epsilon^{(n)}$，即

$$\mathcal{X}^n = \mathcal{T}_\epsilon^{(n)} \bigcup \bar{\mathcal{T}}_\epsilon^{(n)}$$

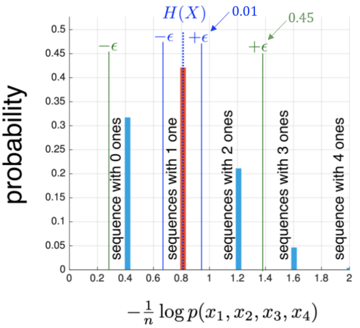
典型集实例

- 考虑独立同分布的二元随机序列 $\boldsymbol{X} = X_1, X_2, X_3, X_4$，且 $p_0 = \frac{3}{4}, p_1 = \frac{1}{4}$。当 $\epsilon = 0.01$ 或 $\epsilon = 0.45$ 时，找出典型集及其元素个数。

典型集实例

- 考虑独立同分布的二元随机序列 $\boldsymbol{X} = X_1, X_2, X_3, X_4$，且 $p_0 = \frac{3}{4}, p_1 = \frac{1}{4}$。当 $\epsilon = 0.01$ 或 $\epsilon = 0.45$ 时，找出典型集及其元素个数。
- 样本熵的分布（可能存在的样本总数为 $2^4 = 16$）

x_1, x_2, x_3, x_4	$p_{\mathbf{X}}(\mathbf{x})$	sample entropy $-\frac{1}{n}\log p_{\mathbf{X}}(\mathbf{x})$
0 0 0 0	0.31641	0.41504
1 0 0 0	0.10547	0.81128
0 1 0 0	0.10547	0.81128
0 0 1 0	0.10547	0.81128
0 0 0 1	0.10547	0.81128
1 1 0 0	0.03516	1.20752
1 0 1 0	0.03516	1.20752
0 1 1 0	0.03516	1.20752
1 0 0 1	0.03516	1.20752
0 1 0 1	0.03516	1.20752
0 0 1 1	0.03516	1.20752



向量信源编码定理

1. 向量信源编码
2. 大数定律
3. 样本熵和典型集
4. **渐近均分性（AEP）**
5. 向量信源编码定理的证明

注：典型集和渐近均分性是信息论的基础

“信息论是大数定律的巧妙应用”

渐近均分性 (Asymptotic Equipartition Property, AEP)

主要性质:

当 $n \to \infty$ 且 $\epsilon \to 0$ 时:

1. **近似等概**: 若 $\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}$, 则 $p_{\boldsymbol{x}} \to 2^{-nH(X)}$, 典型序列近似等概

渐近均分性 (Asymptotic Equipartition Property, AEP)

主要性质:

当 $n \to \infty$ 且 $\epsilon \to 0$ 时:

1. **近似等概**: 若 $\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}$, 则 $p_{\boldsymbol{x}} \to 2^{-nH(X)}$, 典型序列近似等概
2. **高概率**: $\Pr(\boldsymbol{X} \in \mathcal{T}_\epsilon^{(n)}) \to 1$, 典型集的概率很高

渐近均分性 (Asymptotic Equipartition Property, AEP)

主要性质:

当 $n \to \infty$ 且 $\epsilon \to 0$ 时:

1. **近似等概**: 若 $\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}$, 则 $p_{\boldsymbol{x}} \to 2^{-nH(X)}$, 典型序列近似等概
2. **高概率**: $\Pr(\boldsymbol{X} \in \mathcal{T}_\epsilon^{(n)}) \to 1$, 典型集的概率很高
3. **个数少**: $|\mathcal{T}_\epsilon^{(n)}| \to 2^{nH(X)}$, 典型集元素个数很少, $|\mathcal{T}_\epsilon^{(n)}| \ll |\mathcal{X}^n|$

渐近均分性 (Asymptotic Equipartition Property, AEP)

主要性质:

当 $n \to \infty$ 且 $\epsilon \to 0$ 时:

1. **近似等概**: 若 $\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}$, 则 $p_{\boldsymbol{x}} \to 2^{-nH(X)}$, 典型序列近似等概
2. **高概率**: $\Pr(\boldsymbol{X} \in \mathcal{T}_\epsilon^{(n)}) \to 1$, 典型集的概率很高
3. **个数少**: $|\mathcal{T}_\epsilon^{(n)}| \to 2^{nH(X)}$, 典型集元素个数很少, $|\mathcal{T}_\epsilon^{(n)}| \ll |\mathcal{X}^n|$

注: 性质 1 (**近似等概性**) 直接由样本熵和典型集定义可得

$$\bar{H}(\boldsymbol{X}) = -\frac{1}{n}\log p_{\boldsymbol{X}} \xrightarrow[\epsilon \to 0]{n \to \infty} H(X).$$

性质 2（高概率）：随机序列样本大概率是典型的

命题：独立同分布的二元随机序列 $\boldsymbol{X} = X_1, X_2, \cdots, X_n$，其概率分布为 $p(x)$，则 $\forall \epsilon > 0$，有

$$\lim_{n\to\infty} \Pr(\boldsymbol{X} \in \mathcal{T}_\epsilon^{(n)}) = 1$$

性质 2（高概率）：随机序列样本大概率是典型的

命题：独立同分布的二元随机序列 $\boldsymbol{X} = X_1, X_2, \cdots, X_n$，其概率分布为 $p(x)$，则 $\forall \epsilon > 0$，有

$$\lim_{n\to\infty} \Pr(\boldsymbol{X} \in \mathcal{T}_\epsilon^{(n)}) = 1$$

证明：对独立同分布 (IID) 的序列 $\{X_i\}$，

$$\text{IID 随机变量：} \quad Y_i \equiv \bar{H}(X_i) = -\log P_{X_i}, \forall i$$

性质 2（高概率）：随机序列样本大概率是典型的

命题：独立同分布的二元随机序列 $\boldsymbol{X} = X_1, X_2, \cdots, X_n$，其概率分布为 $p(x)$，则 $\forall \epsilon > 0$，有

$$\lim_{n\to\infty} \Pr(\boldsymbol{X} \in \mathcal{T}_\epsilon^{(n)}) = 1$$

证明：对独立同分布 (IID) 的序列 $\{X_i\}$，

$$\text{IID 随机变量：} \quad Y_i \equiv \bar{H}(X_i) = -\log P_{X_i}, \forall i$$

$$\Rightarrow \bar{Y}_n \equiv \frac{1}{n}\sum_{i=1}^{n} Y_i = \bar{H}(\boldsymbol{X})$$

性质 2（高概率）：随机序列样本大概率是典型的

命题：独立同分布的二元随机序列 $\boldsymbol{X} = X_1, X_2, \cdots, X_n$，其概率分布为 $p(x)$，则 $\forall \epsilon > 0$，有

$$\lim_{n \to \infty} \Pr(\boldsymbol{X} \in \mathcal{T}_\epsilon^{(n)}) = 1$$

证明：对独立同分布 (IID) 的序列 $\{X_i\}$，

$$\text{IID 随机变量：}\quad Y_i \equiv \bar{H}(X_i) = -\log P_{X_i}, \forall i$$

$$\Rightarrow \bar{Y}_n \equiv \frac{1}{n}\sum_{i=1}^{n} Y_i = \bar{H}(\boldsymbol{X})$$

$$\Rightarrow \mathrm{E}[Y_i] = \mathrm{E}[Y] = H(X)$$

性质 2（高概率）：随机序列样本大概率是典型的

命题：独立同分布的二元随机序列 $\boldsymbol{X} = X_1, X_2, \cdots, X_n$，其概率分布为 $p(x)$，则 $\forall \epsilon > 0$，有

$$\lim_{n\to\infty} \Pr(\boldsymbol{X} \in \mathcal{T}_\epsilon^{(n)}) = 1$$

证明：对独立同分布 (IID) 的序列 $\{X_i\}$，

$$\text{IID 随机变量：} \quad Y_i \equiv \bar{H}(X_i) = -\log P_{X_i}, \forall i$$

$$\Rightarrow \bar{Y}_n \equiv \frac{1}{n}\sum_{i=1}^{n} Y_i = \bar{H}(\boldsymbol{X})$$

$$\Rightarrow \mathrm{E}[Y_i] = \mathrm{E}[Y] = H(X)$$

由大数定律：$\lim_{n\to\infty} \Pr(|\bar{Y}_n - \mathrm{E}[Y]| \le \epsilon) = 1$，即：

性质 2（高概率）：随机序列样本大概率是典型的

命题：独立同分布的二元随机序列 $\boldsymbol{X} = X_1, X_2, \cdots, X_n$，其概率分布为 $p(x)$，则 $\forall \epsilon > 0$，有

$$\lim_{n\to\infty} \Pr(\boldsymbol{X} \in \mathcal{T}_\epsilon^{(n)}) = 1$$

证明：对独立同分布 (IID) 的序列 $\{X_i\}$，

$$\text{IID 随机变量：} \quad Y_i \equiv \bar{H}(X_i) = -\log P_{X_i}, \forall i$$

$$\Rightarrow \bar{Y}_n \equiv \frac{1}{n}\sum_{i=1}^{n} Y_i = \bar{H}(\boldsymbol{X})$$

$$\Rightarrow \mathrm{E}[Y_i] = \mathrm{E}[Y] = H(X)$$

由大数定律：$\lim_{n\to\infty} \Pr(|\bar{Y}_n - \mathrm{E}[Y]| \leq \epsilon) = 1$，即：

$$\lim_{n\to\infty} \Pr(|\bar{H}(\boldsymbol{X}) - H(X)| < \epsilon) = 1$$

性质 2（高概率）：随机序列样本大概率是典型的

命题：独立同分布的二元随机序列 $\boldsymbol{X} = X_1, X_2, \cdots, X_n$，其概率分布为 $p(x)$，则 $\forall \epsilon > 0$，有

$$\lim_{n \to \infty} \Pr(\boldsymbol{X} \in \mathcal{T}_\epsilon^{(n)}) = 1$$

证明：对独立同分布 (IID) 的序列 $\{X_i\}$，

$$\text{IID 随机变量：} \quad Y_i \equiv \bar{H}(X_i) = -\log P_{X_i}, \forall i$$

$$\Rightarrow \bar{Y}_n \equiv \frac{1}{n} \sum_{i=1}^{n} Y_i = \bar{H}(\boldsymbol{X})$$

$$\Rightarrow \mathrm{E}[Y_i] = \mathrm{E}[Y] = H(X)$$

由大数定律：$\lim_{n \to \infty} \Pr(|\bar{Y}_n - \mathrm{E}[Y]| \le \epsilon) = 1$，即：

$$\lim_{n \to \infty} \Pr(|\bar{H}(\boldsymbol{X}) - H(X)| < \epsilon) = 1 \quad \Rightarrow \quad \lim_{n \to \infty} \Pr(\boldsymbol{X} \in \mathcal{T}_\epsilon^{(n)}) = 1$$

性质 3（个数少）：典型集元素个数很少

命题：当 n 足够大时，典型集 $\mathcal{T}_{\epsilon}^{(n)}$ 的元素个数接近 $2^{nH(X)}$，即：

$$|\mathcal{T}_{\epsilon}^{(n)}| \to 2^{nH(X)}$$

性质 3（个数少）：典型集元素个数很少

命题：当 n 足够大时，典型集 $\mathcal{T}_\epsilon^{(n)}$ 的元素个数接近 $2^{nH(X)}$，即：

$$|\mathcal{T}_\epsilon^{(n)}| \to 2^{nH(X)}$$

证明：当 $n \to \infty$ 时，

$$1 = \sum_{\boldsymbol{x}} p(\boldsymbol{x})$$

性质 3（个数少）：典型集元素个数很少

命题：当 n 足够大时，典型集 $\mathcal{T}_\epsilon^{(n)}$ 的元素个数接近 $2^{nH(X)}$，即：

$$|\mathcal{T}_\epsilon^{(n)}| \to 2^{nH(X)}$$

证明：当 $n \to \infty$ 时，

$$1 = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \to \sum_{\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}} p(\boldsymbol{x})$$

性质 3（个数少）：典型集元素个数很少

命题：当 n 足够大时，典型集 $\mathcal{T}_\epsilon^{(n)}$ 的元素个数接近 $2^{nH(X)}$，即：

$$|\mathcal{T}_\epsilon^{(n)}| \to 2^{nH(X)}$$

证明：当 $n \to \infty$ 时，

$$1 = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \to \sum_{\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}} p(\boldsymbol{x}) \to |\mathcal{T}_\epsilon^{(n)}| 2^{-nH(X)}$$

性质 3（个数少）：典型集元素个数很少

命题：当 n 足够大时，典型集 $\mathcal{T}_\epsilon^{(n)}$ 的元素个数接近 $2^{nH(X)}$，即：

$$|\mathcal{T}_\epsilon^{(n)}| \to 2^{nH(X)}$$

证明：当 $n \to \infty$ 时，

$$1 = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \to \sum_{\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}} p(\boldsymbol{x}) \to |\mathcal{T}_\epsilon^{(n)}| 2^{-nH(X)}$$

即，$|\mathcal{T}_\epsilon^{(n)}| \to 2^{nH(X)}$

性质 3（个数少）：典型集元素个数很少

命题：当 n 足够大时，典型集 $\mathcal{T}_\epsilon^{(n)}$ 的元素个数接近 $2^{nH(X)}$，即：

$$|\mathcal{T}_\epsilon^{(n)}| \to 2^{nH(X)}$$

证明：当 $n \to \infty$ 时，

$$1 = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \to \sum_{\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}} p(\boldsymbol{x}) \to |\mathcal{T}_\epsilon^{(n)}| 2^{-nH(X)}$$

即，$|\mathcal{T}_\epsilon^{(n)}| \to 2^{nH(X)}$

- **当 X 不是均匀分布时，典型集占比趋于 0：**

性质 3（个数少）：典型集元素个数很少

命题：当 n 足够大时，典型集 $\mathcal{T}_\epsilon^{(n)}$ 的元素个数接近 $2^{nH(X)}$，即：

$$|\mathcal{T}_\epsilon^{(n)}| \to 2^{nH(X)}$$

证明：当 $n \to \infty$ 时，

$$1 = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \to \sum_{\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}} p(\boldsymbol{x}) \to |\mathcal{T}_\epsilon^{(n)}| 2^{-nH(X)}$$

即，$|\mathcal{T}_\epsilon^{(n)}| \to 2^{nH(X)}$

- **当 X 不是均匀分布时，典型集占比趋于 0**：

$$\frac{2^{nH(X)}}{2^{n \log |\mathcal{X}|}} = 2^{n(H(X) - \log |\mathcal{X}|)}$$

性质 3（个数少）：典型集元素个数很少

命题：当 n 足够大时，典型集 $\mathcal{T}_\epsilon^{(n)}$ 的元素个数接近 $2^{nH(X)}$，即：

$$|\mathcal{T}_\epsilon^{(n)}| \to 2^{nH(X)}$$

证明：当 $n \to \infty$ 时，

$$1 = \sum_{\boldsymbol{x}} p(\boldsymbol{x}) \to \sum_{\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}} p(\boldsymbol{x}) \to |\mathcal{T}_\epsilon^{(n)}| 2^{-nH(X)}$$

即，$|\mathcal{T}_\epsilon^{(n)}| \to 2^{nH(X)}$

- **当 X 不是均匀分布时，典型集占比趋于 0**：

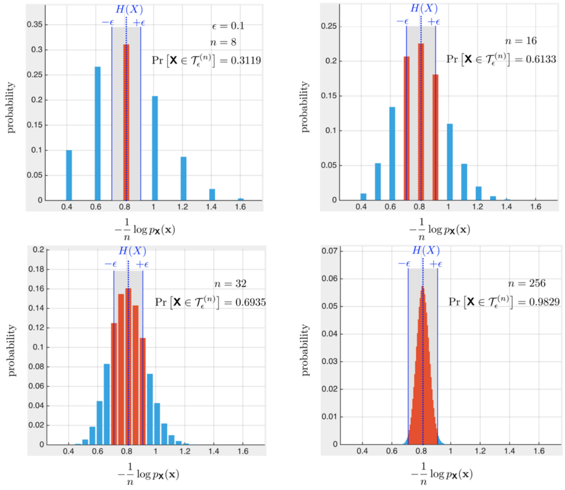
$$\frac{2^{nH(X)}}{2^{n \log |\mathcal{X}|}} = 2^{n(H(X) - \log |\mathcal{X}|)} \xrightarrow[n \to \infty]{H(X) < \log |\mathcal{X}|} 0$$

AEP 实例

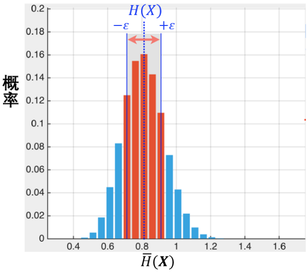
- 考虑独立同分布二元随机变量 $\boldsymbol{X} = X_1, X_2, \cdots, X_n$，且 $p_0 = \frac{3}{4}, p_1 = \frac{1}{4}$。当 $\epsilon = 0.1$ 且 $n = 8, 16, 32, 256$ 时，找出典型集及其概率。

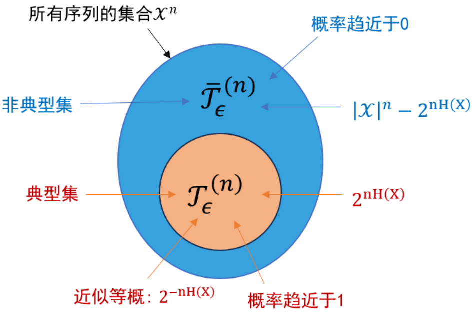
AEP 实例

- 考虑独立同分布二元随机变量 $\boldsymbol{X} = X_1, X_2, \cdots, X_n$，且 $p_0 = \frac{3}{4}, p_1 = \frac{1}{4}$。当 $\epsilon = 0.1$ 且 $n = 8, 16, 32, 256$ 时，找出典型集及其概率。



渐近均分性





渐近均分性：一小部分近似等概的序列占了绝大部分可能性。

向量信源编码定理

1. 向量信源编码
2. 大数定律
3. 样本熵和典型集
4. 渐近均分性（AEP）
5. 向量信源编码定理的证明

注：典型集和渐近均分性是信息论的基础

信源编码定理及其证明

定理 (向量信源编码定理)：令 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)$ 为 n 维独立同分布随机序列，其中 $X_i\sim p(x)$，其信息熵为 $H(X)$。则，对任意 $\epsilon>0$，当 n 足够大时，存在一种码率为 R 的向量编码满足：

$$R\leq H(X)+\epsilon.$$

“存在一种向量编码”：接下来给出一个编码策略

基于典型集的向量编码

令 $\boldsymbol{x} = x_1, x_2, \cdots, x_n$:

- 如果 $\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}$,将 $\mathcal{T}_\epsilon^{(n)}$ 中各个元素与如下集合对应:

$$\{1, 2, 3, \cdots, |\mathcal{T}_\epsilon^{(n)}|\}$$

  码字为 $C(\boldsymbol{x}) =$ "0" + "对应索引"

基于典型集的向量编码

令 $\boldsymbol{x} = x_1, x_2, \cdots, x_n$:

- 如果 $\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}$，将 $\mathcal{T}_\epsilon^{(n)}$ 中各个元素与如下集合对应:

$$\{1, 2, 3, \cdots, |\mathcal{T}_\epsilon^{(n)}|\}$$

  码字为 $C(\boldsymbol{x}) =$ “0” + “对应索引”

- 如果 $\boldsymbol{x} \in \bar{\mathcal{T}}_\epsilon^{(n)}$，则不压缩。码字为 $C(\boldsymbol{x}) =$ “1” + “$\boldsymbol{x}$ 本身”

基于典型集的向量编码

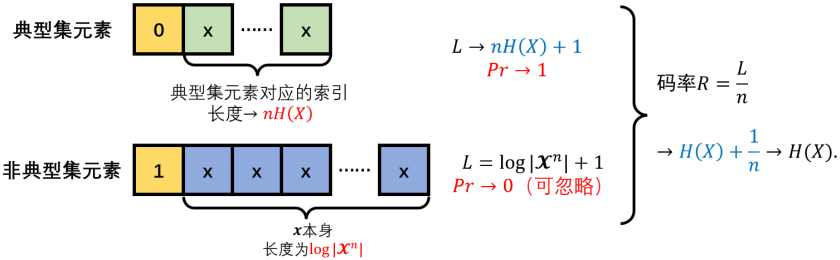
令 $\boldsymbol{x} = x_1, x_2, \cdots, x_n$:

- 如果 $\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}$，将 $\mathcal{T}_\epsilon^{(n)}$ 中各个元素与如下集合对应：

$$\{1, 2, 3, \cdots, |\mathcal{T}_\epsilon^{(n)}|\}$$

  码字为 $C(\boldsymbol{x}) =$ “0” + “对应索引”

- 如果 $\boldsymbol{x} \in \bar{\mathcal{T}}_\epsilon^{(n)}$，则不压缩。码字为 $C(\boldsymbol{x}) =$ “1” + “$\boldsymbol{x}$ 本身”



基于典型集的向量编码

码率计算:

$$\Pr(\boldsymbol{X} \in \bar{\mathcal{T}}_\epsilon^{(n)}) \to 0 \Rightarrow R_{\boldsymbol{x} \in \bar{\mathcal{T}}_\epsilon^{(n)}} \to 0$$

因此,

$$R = R_{\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}} + R_{\boldsymbol{x} \in \bar{\mathcal{T}}_\epsilon^{(n)}}$$

基于典型集的向量编码

码率计算:

$$\Pr(\boldsymbol{X} \in \bar{\mathcal{T}}_\epsilon^{(n)}) \to 0 \Rightarrow R_{\boldsymbol{x}\in\bar{\mathcal{T}}_\epsilon^{(n)}} \to 0$$

因此,

$$\begin{aligned} R &= R_{\boldsymbol{x}\in\mathcal{T}_\epsilon^{(n)}} + R_{\boldsymbol{x}\in\bar{\mathcal{T}}_\epsilon^{(n)}} \\ &\to R_{\boldsymbol{x}\in\mathcal{T}_\epsilon^{(n)}} \qquad \text{``} = R_{\boldsymbol{x}\in\mathcal{T}_\epsilon^{(n)}} + \epsilon'\text{''} \end{aligned}$$

基于典型集的向量编码

码率计算：

$$\Pr(\boldsymbol{X} \in \bar{\mathcal{T}}_\epsilon^{(n)}) \to 0 \Rightarrow R_{\boldsymbol{x} \in \bar{\mathcal{T}}_\epsilon^{(n)}} \to 0$$

因此，

$$\begin{aligned} R &= R_{\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}} + R_{\boldsymbol{x} \in \bar{\mathcal{T}}_\epsilon^{(n)}} \\ &\to R_{\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}} \qquad\qquad\qquad \text{``} = R_{\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}} + \epsilon'\text{''} \\ &= \frac{\log |\mathcal{T}_\epsilon^{(n)}| + 1}{n} \end{aligned}$$

基于典型集的向量编码

码率计算：

$$\Pr(\boldsymbol{X} \in \bar{\mathcal{T}}_\epsilon^{(n)}) \to 0 \Rightarrow R_{\boldsymbol{x} \in \bar{\mathcal{T}}_\epsilon^{(n)}} \to 0$$

因此，

$$\begin{aligned} R &= R_{\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}} + R_{\boldsymbol{x} \in \bar{\mathcal{T}}_\epsilon^{(n)}} \\ &\to R_{\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}} && \text{``} = R_{\boldsymbol{x} \in \mathcal{T}_\epsilon^{(n)}} + \epsilon' \text{''} \\ &= \frac{\log |\mathcal{T}_\epsilon^{(n)}| + 1}{n} \\ &\leq H(X) + \epsilon. && \text{``}\epsilon = \epsilon' + \frac{1}{n}\text{''} \end{aligned}$$

总结

- **向量信源编码**优于**单信源编码**
- **单信源编码**最差可达码率为 $H(X)+1$
- **向量信源编码**最差可达码率为 $H(X)+\epsilon$（当 $n\to\infty$ 时 $\epsilon\to 0$）
- 信息论的核心思路：**大系统**（$n\to\infty$）**下进行证明**
- **典型集**和**渐近均分性**给出了信源编码定理的简易证明

作业

- 复习授课内容
- 预习马尔可夫链和熵率
- 独立完成习题

令 $\boldsymbol{X} = X_1, \cdots, X_n$ 为 IID 随机变量，其分布为 $p_{\boldsymbol{X}}(x_1, \cdots, x_n)$，计算：

$$\lim_{n\to\infty} \left(p_{\mathrm{X}}(x_1, \ldots, x_n)\right)^{1/n}$$