Fundamentals of Information Theory

Homework Three

王翎羽　U202213806　提高 2201 班

2024 年 5 月 29 日

Problem 1 Solutions:

(a) The stationary distribution $\pi = (\pi_0, \pi_1)$ satisfies the equation $\pi P = \pi$ and the condition $\pi_0 + \pi_1 = 1$.

$$\begin{bmatrix} \pi_0 & \pi_1 \end{bmatrix} \begin{bmatrix} 1-p_{01} & p_{01} \\ p_{10} & 1-p_{10} \end{bmatrix} = \begin{bmatrix} \pi_0 & \pi_1 \end{bmatrix}$$

Thus, $\pi_0 = \frac{p_{10}}{p_{10}+p_{01}}$ and $\pi_1 = \frac{p_{01}}{p_{10}+p_{01}}$. By $H = -\sum_i \pi_i \sum_j P_{ij} \log P_{ij}$, we have

$$H(\mathcal{X}) = -\left[\frac{p_{10}\log(1-p_{01}) + p_{01}\log(1-p_{10}) + p_{10}p_{01}(\log p_{01} + \log p_{10})}{p_{01}+p_{10}}\right]$$

(b) The entropy rate is maximized when the transitions are as uncertain as possible, which corresponds to $p_{01} = p_{10} = 0.5$. So $H(\mathcal{X}) = 1$.

(c) In the former question, let $p = p_{01}$, $1 = p_{10}$, we can get this situation.
So it is $H(\mathcal{X}) = -\frac{1}{1+p}\left[(1-p)\log(1-p) + p\log p\right]$.

(d) From Lagrange multiplier method,

$$\max_{H(\mathcal{X})} \quad L(p,\lambda) = H(\mathcal{X}) + \lambda \cdot (0.5 - p), \quad p \le 0.5$$

So $\lambda = \frac{2}{3}$, and $H(\mathcal{X})_{max} = 1$.

Problem 2 Solutions:

(a) Not necessarily. For compact code, the overhead is at most 1 bit, but for shanon code, let me explain with an extreme example, if we let one of these probabilities approach infinitesimal, then you will see the overhead goes to infinity.

(b) Look at my answer to Problem 3, they are the same.

Problem 3 Solutions:

(a) $l_i = \lceil \log \frac{1}{p(x_i)} \rceil$,

$$H(x) = \sum_i p(x) \log \frac{1}{p(x_i)} < L = \sum_i p(x) l_i = \sum_i p(x) \lceil \log \frac{1}{p(x_i)} \rceil < H(x) + 1$$

Assuming that x_p is a prefix of x_{p+q}, so we have $F_{p+q} - F_p < 2^{-l_p}$ because they are the same in the former p bits. And then $\sum_{p}^{p+q-1} p_i < 2^{-l_p}$, obviously, $p_k < 2^{-l_p}$. However, $l_p = \lceil \log \frac{1}{p(x_p)} \rceil \ge \log \frac{1}{p(x_p)}$, so $2^{-l_p} \ge p_k$.So this contradicts our hypothesis, therefore the Shannon code is prefix-code.

(b)

i	p_i	$l_i = \lceil \log \frac{1}{p(x_i)} \rceil$	F_i	Codeword
1	0.5	1	0	0
2	0.25	2	0.5	10
3	0.125	3	0.75	110
4	0.125	3	0.875	111

Problem 4 Solutions:

(a)

X	1	2	3	4	5	6	7
Codeword	0	10	110	1110	11110	111110	111111

(b)

$$L = \sum_{i=0}^{6} p_i \cdot l_i = 0.49 \times 1 + 0.26 \times 2 + 0.12 \times 3 + 0.04 \times 4 + 0.04 \times 5 + 0.03 \times 6 + 0.02 \times 6 = 2.03$$

(c)

X	1	2	3	4	5	6	7
Codeword	0	1	20	22	212	211	210

Problem 5 Solutions:

(a)

$$\sum_{i=1}^{6} p_i \cdot l_i = 1 \times \frac{8}{23} + 2 \times \frac{6}{23} + 3 \times \frac{4}{23} + 4 \times \frac{2}{23} + 5 \times \frac{2}{23} + 5 \times \frac{1}{23} = \frac{55}{23} \approx 2.39$$

(b)The one with probability $\frac{8}{23}$.

(c) Huffman coding.

$$\sum_{i=1}^{6} p_i \cdot l_i = 2 \times \frac{8}{23} + 2 \times \frac{6}{23} + 2 \times \frac{4}{23} + 3 \times \frac{2}{23} + 4 \times \frac{2}{23} + 4 \times \frac{1}{23} = \frac{55}{23} \approx 2.35$$

(d) Mixture of the first two bottles.