信息论基础（于秀兰 陈前斌 王永）课后作业答案

注：X 为随机变量，概率$P(X=x)$是 x 的函数，所以$P(X)$仍为关于 X 的随机变量，文中如无特别说明，则以此类推。

第一章

1.6

$$[P(xy)]=\begin{bmatrix}P(b_1a_1) & P(b_2a_1)\\ P(b_1a_2) & P(b_2a_2)\end{bmatrix}=\begin{bmatrix}0.36 & 0.04\\ 0.12 & 0.48\end{bmatrix}$$

$$[P(y)]=[P(b_1)\quad P(b_2)]=[0.48\quad 0.52]$$

$$[P(x|y)]=\begin{bmatrix}P(a_1|b_1) & P(a_2|b_1)\\ P(a_1|b_2) & P(a_2|b_2)\end{bmatrix}=\begin{bmatrix}0.75 & 0.25\\ 0.077 & 0.923\end{bmatrix}$$

第二章

2.1

(1)

$$I(B)=-\log P(B)=-\log\frac{1}{8}=3(\text{bit})$$

注：此处$P(B)$表示事件 B 的概率。

(2)

设信源为 X，

$$H(X)=E[-\log P(X)]=-\frac{1}{4}log\frac{1}{4}-2\cdot\frac{1}{8}log\frac{1}{8}-\frac{1}{2}log\frac{1}{2}=1.75(bit/symbol)$$

(3)

$$\xi=1-\eta=1-\frac{1.75}{log4}=12.5\%$$

2.2

(1)

P(3 和 5 同时出现)=1/18

$$I=-\log\frac{1}{18}\approx 4.17(\text{bit})$$

(2)

P(两个 2 同时出现)=1/36

$$I=-\log\frac{1}{36}\approx 5.17(\text{bit})$$

(3)

向上点数和为 5 时（14，23，41，32）有 4 种，概率为 1/9，

$\mathrm{I}=-\log\frac{1}{9}\approx 3.17(\mathrm{bit})$

(4)

设两个点数和为 X，则

X	2	3	4	5	6	7	8	9	10	11	12
P(X)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

$\mathrm{H(X)=E[-logP(X)]}\approx 3.27(bit/symbol)$

(5)

$\mathrm{P}(\text{两个点数至少有一个 }1)=1-\frac{5}{6}\cdot\frac{5}{6}=\frac{11}{36}$

$\mathrm{I}=-\log\frac{11}{36}\approx 1.71(\mathrm{bit})$

(6)

相同点数有 6 种，概率分别为 1/36；

不同点数出现有 15 种，概率分别为 1/18；

$\mathrm{H}=6\cdot\frac{1}{36}\cdot\log 36+15\cdot\frac{1}{18}\cdot\log 18\approx 4.34(\mathrm{bit/symbol})$

2.9

(1)

$$\mathrm{H(X,Y)=E[-logP(X,Y)]}=-\sum_{i=1}^{3}\sum_{j=1}^{3}P\left(x_i,y_j\right)logP\left(x_i,y_j\right)\approx 2.3(bit/sequence)$$

(2)

$\mathrm{H(Y)=E[-logP(Y)]}\approx 1.59(\mathrm{bit/symbol})$

(3)

$\mathrm{H(X|Y)=H(X,Y)-H(Y)}=0.71(\mathrm{bit/symbol})$

2.12

(1)

$\mathrm{H(X)=E[-logP(X)]}=-\frac{2}{3}log\frac{2}{3}-\frac{1}{3}log\frac{1}{3}\approx 0.92(\mathrm{bit/symbol})$

Y 的分布律为：$1/2, 1/3, 1/6$；

$\mathrm{H(Y)=E[-logP(Y)]}\approx 1.46(\mathrm{bit/symbol})$

(2)

$$\mathrm{H}(\mathrm{Y}|a_1)=\mathrm{E}[-\mathrm{logP(Y|X)}|X=a_1]=-\sum_i P(b_i|a_1)logP(b_i|a_1)$$

$$=-\frac{3}{4}log\frac{3}{4}-\frac{1}{4}log\frac{1}{4}\approx 0.81(bit/symbol)$$

$$\mathrm{H}(\mathrm{Y}|a_2)=\mathrm{E}[-\mathrm{logP(Y|X)}|X=a_2]=-\sum_i P(b_i|a_2)logP(b_i|a_2)$$

$$= -\frac{1}{2}log\frac{1}{2} - \frac{1}{2}log\frac{1}{2} = 1(bit/symbol)$$

(3)

$$\mathrm{H(Y|X)} = \sum_i P(a_i)\mathrm{H}(\mathrm{Y}|a_i) = \frac{2}{3}\cdot 0.81 + \frac{1}{3}\cdot 1 \approx 0.87(bit/symbol)$$

2.13

(1)

$\mathrm{H(X)} = \mathrm{H}(0.3, 0.7) \approx 0.88(bit/symbol)$

二次扩展信源的数学模型为随机矢量$X^2 = (X_1X_2)$，其中X_1、X_2和 X 同分布，且相互独立，则

$\mathrm{H}(X^2) = 2\mathrm{H(X)} = 1.76(bit/sequence)$

平均符号熵

$H_2(X^2) = \mathrm{H(X)} \approx 0.88(bit/symbol)$

(2)

二次扩展信源的数学模型为随机矢量$X^2 = (X_1X_2)$，其中X_1、X_2和 X 同分布，且X_1、X_2相关，

$$\mathrm{H}(X_2|X_1) = \mathrm{E}[-\mathrm{logP}(X_2|X_1)] = -\sum_{x_1}\sum_{x_2} P(x_1, x_2)\mathrm{logP}(x_2|x_1)$$

$$= -\frac{1}{10}log\frac{1}{3} - \frac{2}{10}log\frac{2}{3} - \frac{21}{40}log\frac{3}{4} - \frac{7}{40}log\frac{1}{4} \approx 0.84(\mathrm{bit/symbol})$$

$\mathrm{H}(X^2) = \mathrm{H}(X_1, X_2) = \mathrm{H}(X_2|X_1) + H(X_1) = 0.84 + 0.88 = 1.72(bit/sequence)$

$H_2(X^2) = \mathrm{H}(X^2)/2 = 0.86(bit/symbol)$

2.14

(1)

令无记忆信源为 X，

$$\mathrm{H(X)} = \mathrm{H}\left(\frac{1}{4}, \frac{3}{4}\right) = \frac{1}{4}\times 2 + \frac{3}{4}\times 0.415 \approx 0.81(bit/symbol)$$

(2)

$$\mathrm{I}(X^{100}) = -\mathrm{logP}(X^{100} = x_1x_2 \dots x_{100}) = -log\left[\left(\frac{1}{4}\right)^m\left(\frac{3}{4}\right)^{100-m}\right]$$

$$= 2m + (2 - log3)(100 - m) = 200 - (100 - m)log3\ (bit)$$

(3)

$\mathrm{H}(X^{100}) = 100H(X) = 81(bit/sequence)$

2.15

(1)

因为信源序列符号间相互独立，且同分布，所以信源为一维离散平稳信源。

(2)

$\mathrm{H(X)} = \mathrm{H}(0.2, 0.8) \approx 0.72(bit/symbol)$
$\mathrm{H}(X^2) = 2H(X) = 1.44(bit/sequence)$
$\mathrm{H}(X_3|X_1X_2) = H(X_3) = H(X) = 0.72(bit/symbol)$
$H_\infty = H(X) = 0.72(bit/symbol)$

2.16

(1)

$$\mathrm{H}(X_2|X_1) = \mathrm{E}[-\mathrm{logP}(X_2|X_1)] = -\sum_{x_1}\sum_{x_2} P(x_1, x_2)\mathrm{logP}(x_2|x_1)$$

$$= -\frac{6}{10}log\frac{9}{10} - \frac{2}{30}log\frac{1}{10} - \frac{2}{30}log\frac{2}{10} - \frac{8}{30}log\frac{8}{10} \approx 0.55(\mathrm{bit/symbol})$$

$\mathrm{H}(X_3|X_2X_1) = \ \mathrm{H}(X_3|X_2) = \mathrm{H}(X_2|X_1) \approx 0.55(\mathrm{bit/symbol})$
$\mathrm{H}(X_4|X_3X_2X_1) = \ \mathrm{H}(X_4|X_3) = \mathrm{H}(X_2|X_1) \approx 0.55(\mathrm{bit/symbol})$

(2)

$H_\infty = \mathrm{H}(X_2|X_1) \approx 0.55(\mathrm{bit/symbol})$

$$\xi = 1 - \eta = 1 - \frac{0.55}{log2} = 45\%$$

(3)

$$\mathrm{H(X)} = \mathrm{H}\left(\frac{2}{3}, \frac{1}{3}\right) \approx 0.92(bit/symbol)$$

$H_\infty \leq H(X)$，二维离散平稳信源的极限熵不大于其单符号信源的熵，说明离散单符号信源扩展后的单符号平均熵是非增的。

2.18

(1)

$a_i \epsilon A$，A 是状态集；$\mathrm{P}(x_{i+1}|s_i = E_a)$表示 i 时刻状态为$E_a$，$\mathrm{i} + 1$时刻输出$x_{i+1}$。
该马尔科夫链的状态转移矩阵为

$$\mathrm{P} = \left[\mathrm{P}\left(E_{a_j}\middle|E_{a_i}\right)\right] = [\mathrm{P}(x_{i+1}|x_i)] = [\mathrm{P}(x_{i+1}|s_i = E_a)] = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 2/3 & 0 & 1/3 \\ 2/3 & 1/3 & 0 \end{bmatrix},$$

$$P^2 = \begin{bmatrix} 7/12 & 5/24 & 5/24 \\ 5/9 & 5/18 & 1/6 \\ 5/9 & 1/6 & 5/18 \end{bmatrix},$$

所以该链为齐次遍历马尔科夫链。

(2)

令$\mathrm{P}(x_i = k) = p_i(\mathrm{k})$，则
$[p_1(1) \quad p_1(2) \quad p_1(3)] = [1/2 \quad 1/4 \quad 1/4]$，
$[p_2(1) \quad p_2(2) \quad p_2(3)] = [p_1(1) \quad p_1(2) \quad p_1(3)]P$

$$= \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{bmatrix} = \begin{bmatrix} \frac{7}{12} & \frac{5}{24} & \frac{5}{24} \end{bmatrix}$$

因为$[p_1(1) \quad p_1(2) \quad p_1(3)] \neq [p_2(1) \quad p_2(2) \quad p_2(3)]$，所以该信源不是离散平稳信源。

(3)

当信源的输出序列足够长，马尔科夫链达到平稳分布时，该信源可以看作离散平稳信源。

(4)

$$\mathrm{H}(X_{i+1}|s_i = E_1) = -\sum_a P(x_{i+1} = a|s_i = E_1) log P(x_{i+1} = a|s_i = E_1)$$

$$= -\frac{1}{2}log\frac{1}{2} - 2\cdot\frac{1}{4}log\frac{1}{4} = 1.5(bit/symbol)$$

同理得：

$\mathrm{H}(X_{i+1}|s_i = E_2) \approx 0.92(bit/symbol)$

$\mathrm{H}(X_{i+1}|s_i = E_3) \approx 0.92(bit/symbol)$

设极限分布为$[\mathrm{P}(E_1) \quad \mathrm{P}(E_2) \quad \mathrm{P}(E_3)]$，则

$$\mathrm{P}(E_1) = \frac{1}{2}\mathrm{P}(E_1) + \frac{2}{3}\mathrm{P}(E_2) + \frac{2}{3}\mathrm{P}(E_3)$$

$$\mathrm{P}(E_2) = \frac{1}{4}\mathrm{P}(E_1) + \frac{1}{3}\mathrm{P}(E_3)$$

$$\mathrm{P}(E_3) = \frac{1}{4}\mathrm{P}(E_1) + \frac{1}{3}\mathrm{P}(E_2)$$

$$\mathrm{P}(E_1) + \mathrm{P}(E_2) + \mathrm{P}(E_3) = 1$$

解得

$\mathrm{P}(E_1) = 4/7$，$\mathrm{P}(E_2) = 3/14$，$\mathrm{P}(E_3) = 3/14$

$$H_\infty = H(X_{i+1}|s_i) = \frac{4}{7}\times 1.5 + 2\times\frac{3}{14}\times 0.92 \approx 1.25(bit/symbol)$$

(5)

$H_0 = log3 \approx 1.59(bit/symbol)$

$$H_1 = \mathrm{H}\left(\frac{1}{2},\frac{1}{4},\frac{1}{4}\right) = 1.5(bit/symbol)$$

$$\xi_1 = 1 - \frac{1.5}{1.59} \approx 5.66\%$$

$$H_2 = H(\frac{7}{12} \quad \frac{5}{24} \quad \frac{5}{24}) \approx 1.40(bit/symbol)$$

$$\xi_2 = 1 - \frac{1.4}{1.59} \approx 11.95\%$$

$[p_3(1) \quad p_3(2) \quad p_3(3)] = [p_2(1) \quad p_2(2) \quad p_2(3)]P$

$$= \begin{bmatrix}\frac{7}{12} & \frac{5}{24} & \frac{5}{24}\end{bmatrix}\begin{bmatrix}\frac{1}{2} & \frac{1}{4} & \frac{1}{4}\\ \frac{2}{3} & 0 & \frac{1}{3}\\ \frac{2}{3} & \frac{1}{3} & 0\end{bmatrix} = \begin{bmatrix}\frac{41}{72} & \frac{31}{144} & \frac{31}{144}\end{bmatrix}$$

$H_2 = H(\frac{41}{72} \quad \frac{31}{144} \quad \frac{31}{144}) \approx 1.42(bit/symbol)$

$\xi_3 = 1 - \frac{1.42}{1.59} \approx 10.69\%$

2.20

(1)

状态转移矩阵

$$\mathrm{P} = [\mathrm{P}(E_j|E_i)] = \begin{bmatrix} 0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 \end{bmatrix}$$

(2)

由 P 知此马尔科夫链存在极限分布，

设极限分布为$[\mathrm{P}(E_1) \quad \mathrm{P}(E_2) \quad \mathrm{P}(E_3) \quad \mathrm{P}(E_4)]$，则

$\mathrm{P}(E_1) = 0.8\mathrm{P}(E_1) + 0.5\mathrm{P}(E_3)$

$\mathrm{P}(E_2) = 0.2\mathrm{P}(E_1) + 0.5\mathrm{P}(E_3)$

$\mathrm{P}(E_3) = 0.5\mathrm{P}(E_2) + 0.2\mathrm{P}(E_4)$

$\mathrm{P}(E_4) = 0.5\mathrm{P}(E_2) + 0.8\mathrm{P}(E_4)$

$\mathrm{P}(E_1) + \mathrm{P}(E_2) + \mathrm{P}(E_3) + \mathrm{P}(E_4) = 1$

解得

$\mathrm{P}(E_1) = 5/14$，$\mathrm{P}(E_2) = 1/7$，$\mathrm{P}(E_3) = 1/7$，$\mathrm{P}(E_4) = 5/14$

(3)

$\mathrm{H}(X_{i+1}|s_i = E_4) = \mathrm{H}(X_{i+1}|s_i = E_1) = -0.8log0.8 - 0.2log0.2$
$\approx 0.72(bit/symbol)$

$\mathrm{H}(X_{i+1}|s_i = E_3) = \mathrm{H}(X_{i+1}|s_i = E_2) = -2 * 0.5log0.5 = 1(bit/symbol)$

$H_\infty = H(X_{i+1}|s_i) = 2 \times \frac{5}{14} \times 0.72 + 2 \times \frac{1}{7} \times 1 = 0.8(bit/symbol)$

(4)

$$\mathrm{P}(0) = \sum_i P(0|E_i)P(E_i) = 0.8 \times \frac{5}{14} + 0.5 \times \frac{1}{7} + 0.5 \times \frac{1}{7} + 0.2 \times \frac{5}{14} = 0.5$$

$\mathrm{P}(1) = 0.5$

(5)

初始时刻的$\mathrm{P}(0), \mathrm{P}(1)$和(4)中不一样，所以初始时刻的信源不是平稳信源。