第二章习题解答分为两部分：

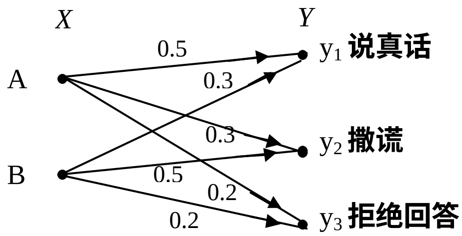
PartA(1-11 页)：1，2，3，4，5，8，9，18，19，22，24，25，

PartB(12-17 页)：5（另解），6，10，11，12，20

《信息论与编码》第二章习题解答

2.1 A **村有一半人说真话，$\frac{3}{10}$人总说假话，$\frac{2}{10}$人拒绝回答；**B **村有**$\frac{3}{10}$**人诚实，一半人说谎，**$\frac{2}{10}$**人拒绝回答。现随机地从** A **村和** B **村抽取人，**p **为抽到** A **村人的概率，**$1-p$ **为抽到** B **村人的概率，问通过测试某人说话的状态平均能获得多少关于该人属于哪个村的信息？通过改变** p**，求出该信息的最大值。**

[解] **用** X **表示随机抽取人所属的村别，**Y **表示说话的状态，则** X **和** Y **之间的关系图如下所示。**



$$P(A)+P(B)=1$$

$$P(y_1)=0.5\cdot p(A)+0.3[1-p(A)]$$

$$P(y_2)=0.3\cdot p(A)+0.5[1-p(A)]$$

$$P(y_3)=0.2\,p(A)+0.2[1-p(A)]=0.2$$

$$H(Y)=-\sum_{i=1}^{3}p(y_i)\log p(y_i)$$

$$H(Y|X)=P(A)H(Y|X=A)+P(B)H(Y|X=B)$$

$$=H(0.5,0.3,0.2)$$

$$=-0.5\log 0.5-0.3\log 0.3-0.2\log 0.2=1.485\ \text{bit}$$

$$I(X;Y)=H(Y)-H(Y|X)$$

由对称性，当 $P(A)=P(B)=0.5$ **时，互信息最大，这时**

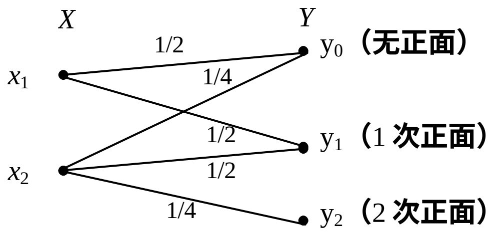
$$H(Y)=0.4\log 0.4+0.4\log 0.4+0.2\log 0.2=1.522\ \text{bit}$$

所以

$$I(X;Y)=H(Y)-H(Y|X)=0.037\ \text{bit}$$

2.2 一个无偏骰子，抛掷一次，如果出现 1**，**2**，**3**，**4 **点，则把一枚均匀硬币投掷一次，如果骰子出现** 5**，**6 **点，则硬币投掷二次，求硬币投掷中正面出现次数对于骰子出现点数所提供的信息？**

［解］ **令** $X=x_1$ **表示掷骰子出现 1，2，3，4 点，** $X=x_2$ **表示出现 5，6 点，** Y **表示出现硬币正面的次数，于是** X **和** Y **具有如下关系图。**



$$P(X=x_1)=2/3,\quad P(X=x_2)=1/3$$

$$P(Y=y_0)=5/12,\quad p(Y=y_1)=1/2,\quad p(Y=y_2)=1/2$$

所以

$$I(X;Y)=H(Y)-H(Y|X)$$

$$=H(\frac{5}{12},\frac{1}{2},\frac{1}{12})-\frac{2}{3}\cdot H(Y|X=x_1)-\frac{1}{3}H(Y|X=x_2)$$

$$=H(\frac{5}{12},\frac{1}{2},\frac{1}{12})-\frac{2}{3}\cdot H(\frac{1}{2},\frac{1}{2})-\frac{1}{3}\cdot H(\frac{1}{4},\frac{1}{2},\frac{1}{4})$$

=0.158 bit

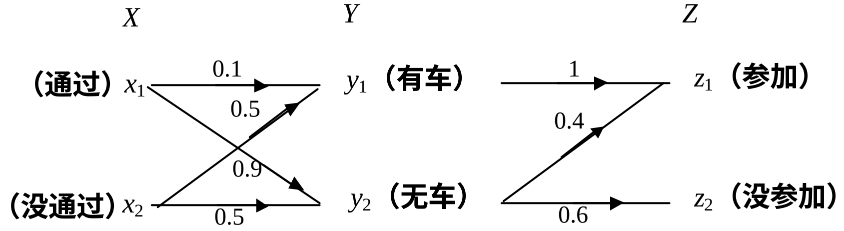
2.3 在某中学有 $\frac{3}{4}$ **学生通过了考试，** $\frac{1}{4}$ **学生没有通过。在通过考试的同学中 10%有自行车，而没有通过的学生中 50%有自行车，所有有自行车的同学都加入了联谊会，无自行车的同学中仅有 40%加入联谊会。**

a. **通过询问是否有自行车，能获得多少关于学生考试成绩的信息？**

b. **通过询问是否参加联谊会，能获得多少关于学生成绩的信息？**

c. **如果把学生成绩情况，自行车拥有情况和是否参加联谊会用三位二进数字传输，问每位数字携带多少信息？**

［解］ X **表示学生有无通过考试，** Y **表示学生有无自行车，** Z **表示学生有无参加联谊会，** X**，**Y**，**Z **之间的关系图。**



$P(x_1)=0.75$， $P(x_2)=0.25$

$P(y_1)=0.2$， $P(y_2)=0.8$

$P(z_1)=0.52$， $P(z_2)=0.48$

$P(z_1|x_1)=0.46$， $P(z_2|x_1)=0.54$

$P(z_1|x_2)=0.7$， $P(z_2|x_2)=0.3$

（a） $I(X;Y)=H(Y)-H(Y|X)$

$$=H(0.2,0.8)-0.75H(0.1,0.9)-0.25H(0.5,0.5)$$

$$=0.12 \text{ bit}$$

（b） $I(X;Z)=H(Z)-H(Z|X)$

$$=H(0.52,0.48)-0.75H(0.46,0.54)-0.25H(0.7,0.3)$$

$$=0.03 \text{ bit}$$

（c）**第一位数字携带信息为** $H(X)=H(0.75,0.25)=0.811 \text{ bit}$

在已知第一位数字下，第二位数字携带信息为

$$H(Y|X)=0.75H(0.1,0.9)+0.25H(0.5,0.5)$$

=0.602 bit

在已知前二位数字下，第三位数字携带信息为：

$H(Z|X,Y)=H(Z|Y)$ **（因为** $X\to Y\to Z$**）**

$$=0.2H(1)+0.8H(0.4,0.6)$$

$$=0.8H(0.4,0.6)$$

$$=0.777 \text{ bit}$$

2.4 随机掷三颗骰子，以 X **表示第一颗骰子抛掷的结果，以** Y **表示第一颗和第二颗骰子抛掷之和，以** Z **表示三颗骰子的点数之和，试求** $H(X|Y)$，$H(Y|X)$，$H(Z|X,Y)$，$H(X,Z|Y)$**和** $H(Z|X)$。

[解] 设第一颗骰子结果为 X_1，**第二颗骰子结果为** X_2，**第二颗结果为** X_3，**则** X_1，X_2，X_3 **是独立同分布随机变量** $X=X_1$， $Y=X_1+X_2$， $Z=X_1+X_2+X_3$

X **和** Y **的事件空间和对应概率为**

X	1	2	3	4	5	6
$p(x)$	1/6	1/6	1/6	1/6	1/6	1/6

Y	2	3	4	5	6	7	8	9	10	11	12
$p(y)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

所以

$$H(X)=H(X_1)=H(X_2)=H(X_3)=\log_2 6=2.585\ \text{bit}$$

$$H(Y)=H(X_1+X_2)=-\sum_{y_i}P(y_i)\log P(y_i)$$

$$=3.274\ \text{bit}$$

$$H(Z|Y)=H(X_3)=\ H(X)=\ 2.585\ \text{bit}$$

$$H(X|Y)=H(X)+H(Y|X)-H(Y)$$

$$=H(X)+H(X_2)-H(Y)$$

$$=2H(X)-H(Y)$$

$$=1.896\ \text{bit}$$

$$H(Y|X)=H(X_2)=H(X)=2.585\ \text{bit}$$

$$H(Z|X,Y)=H(X_3)=H(X)=2.585\ \text{bit}$$

$$H(Z|X)=H(X_2+X_3)=H(Y)=3.274\ \text{bit}$$

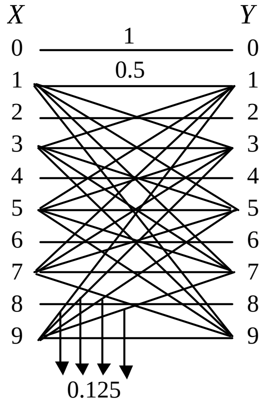
$$H(X,Z|Y)=H(X|Y)+H(Z|X,Y)$$

$$=1.896+2.585$$

$$=4.481\ \text{bit}$$

2.5　设一个系统传送 10 个数字：0，1，2，…，9，奇数在传送时以 0.5 概率等可能地错成另外的奇数，而其他数字总能正确接收。试求收到一个数字后平均得到的信息量。

[解] X **表示发送数字，**Y **表示接收到数字。**



$$I(X;Y)=\sum_x\sum_y p(x,y)\log\frac{p(x|y)}{p(x)}$$

在上式求和中，使 $p(x,y)\neq 0$ **的输入，输出对（**x, y**）可分为 3 类：**

$$S_1=\{(0,0),(2,2),(4,4),(6,6),(8,8)\}$$

$$S_2=\{(1,1),(3,3),(5,5),(7,7),(9,9)\}$$

$$S_3 = \begin{Bmatrix} (1,3),(1,5),(1,7),(1,9) \\ (3,1),(3,5),(3,7),(3,9) \\ (5,1),(5,3),(5,7),(5,9) \\ (7,1),(7,3),(7,5),(7,9) \\ (9,1),(9,3),(9,5),(9,7) \end{Bmatrix}$$

$$\sum_{(x,y)\in S_1} p(x,y)\log\frac{p(x|y)}{p(x)} = 5.p(x=0,y=0)\log\frac{(X=0|Y=0)}{p(X=0)}$$

$$= 5\cdot 0.1\cdot \log\frac{1}{0.1} = 0.5\log 10 = 1.6609$$

$$\sum_{(x,y)\in S_2} p(x,y)\log\frac{p(x|y)}{p(x)} = 5\cdot 0.05\cdot\log\frac{0.5}{0.1} = 0.25\log 5 = 0.5805$$

$$\sum_{(x,y)\in S_3} p(x,y)\log\frac{p(x,y)}{p(x)} = 20\cdot p(X=1,Y=3)\log\frac{p(X=1|Y=3)}{p(X=1)}$$

$$= 20\cdot 0.0125\log\frac{0.125}{0.1} = 0.08$$

所以 $I(X;Y) = 2.3214$ bit

2.8 定义 $S(X,Y) = 1 - r(X,Y) = I(X;Y)/H(X,Y)$ **为随机变量** X **和** Y **之间的相似度，证明**

(a) $0 \le S(X, Y) \le 1$

(b) $S(X, X)=1$

(c) **当** X **和** Y **统计独立时，**$S(X, Y)=0$

[证明]

(a) **因为** $I(X;Y)\ge 0, H(X,Y)\ge 0$

所以 $S(X,Y) = \dfrac{I(X;Y)}{H(X,Y)} \ge 0$ **；**

又 $S(X,Y) = \dfrac{I(X;Y)}{H(X,Y)}$

$$= \frac{H(X)-H(X|Y)}{H(X)+H(Y|X)} \le 1$$ **；**

(b) $S(X,X) = \dfrac{I(X;X)}{H(X,X)} = \dfrac{H(X)}{H(X)} = 1$

(c) **若** X **和** Y **统计独立，则** $H(X|Y) = H(X)$**，**

所以 $S(X,Y) = \dfrac{I(X;Y)}{H(X,Y)} = \dfrac{0}{H(X,Y)} = 0$

2.9 若三个随机变量 X，Y，Z，有 $X+Y=Z$ 成立，其中 X 和 Y 独立，试证

(a) $H(X) \leq H(Z)$

(b) $H(Y) \leq H(Z)$

(c) $H(X,Y) \geq H(Z)$

(d) $I(X;Z)=H(Z)-H(Y)$

(e) $I(X,Y;Z)=H(Z)$

(f) $I(X;Y,Z)=H(X)$

(g) $I(Y;Z|X)=H(Y)$

(h) $I(X;Y|Z)=H(X|Z)=H(Y|Z)$

[证明]

因为 $X+Y=Z$

所以 $p\{Z=z_k \mid X=x_i\}=p\{Y=z_k-x_i \mid X=x_i\}$

所以对给定 $X=x_i$ **有**

$$H(Z \mid X=x_i)=H(Y \mid X=x_i)$$

所以 $H(Z \mid X)=H(Y \mid X)$

因为 X,Y **独立，所以**

$$H(X,Y,Z)=H(Z)+H(X,Y \mid Z)=H(X,Y)+H(Z \mid X,Y)$$
$$=H(X,Y)$$

(c) $H(X,Y)=H(X)+H(Y)\geq H(Z)$

(b) $H(Z)=H(X)+H(Y)-H(X,Y \mid Z)$

$$=H(X)+H(Y)-H(X \mid Z)-H(Y \mid Z,X)$$
$$=H(X)+H(Y)-H(X \mid Z)$$
$$=H(Y)+I(X;Z)$$

所以 $H(Z)\geq H(Y)$ **，同理**

(a) $H(Z)\geq H(X)$

(d) $H(Z)-H(Y)=I(X;Z)$

(e) $I(XY;Z)=H(Z)-H(Z \mid XY)=H(Z)$

(f) $I(X;YZ)=H(X)-H(X \mid YZ)=H(X)$

(g) $I(Y;Z \mid X)=H(Y \mid X)-H(Y \mid Z,X)=H(Y)$

(h) $I(X;Y \mid Z)=H(X \mid Z)-H(X \mid Z,Y)=H(X \mid Z)$

$$=H(Y \mid Z)$$

2.18 令 U 是非负整数集合，事件 $k\in U$ 的概率为 $p(k)$，且 $\sum_{k=0}^{\infty} kp(k)=A$，试求 H（U）达到最大的分布 $\{p(k)\}$。

［解］利用拉格朗日乘子法，引入参数 l_1，和 l_2，求如下目标函数的最大值

$$J(\{p_k\})=-\sum_{k=0}^{\infty} p(k)\ln p(k)-l_1\sum_{k=0}^{\infty} p(k)-l_2\sum_{k=0}^{\infty} kp(k)$$

其中 l_1 和 l_2 由约束条件

$$\sum_{k=0}^{\infty} p(k)=1$$

和

$$\sum_{k=0}^{\infty} k\,p(k)=A$$

来待定。

$$J(\{p_k\})=\sum_{k=0}^{\infty} p(k)\ln\frac{e^{-l_1-l_2k}}{p(k)}$$

$$\le\sum_{k=0}^{\infty} p(k)[\frac{e^{-l_1-l_2k}}{p(k)}-1]$$

当 $\frac{e^{-l_1-l_2k}}{p(k)}=1$ 时上式等号成立，也就达到目标函数 J 的最大值。这时

$$p(k)=e^{-l_1-l_2k}\text{，}\quad k=0,1,2,\cdots$$

由约束条件，

$$\sum_{k=0}^{\infty} p(k)=e^{-l_1}\sum_{k=0}^{\infty} e^{-l_2k}=1$$

$$\sum_{k=0}^{\infty} kp(k)=e^{-l_1}\sum_{k=0}^{\infty} ke^{-l_2k}=A$$

设 $A_1=e^{-l_1},A_2=e^{-l_2}$，即

$$\begin{cases}\dfrac{A_1}{1-A_2}=1\\[2ex]\dfrac{A_1\cdot A_2}{(1-A_2)^2}=A\end{cases}$$

解出 $A_2=\dfrac{A}{1+A},A_1=\dfrac{1}{1+A}$

所以达到最大熵的分布为

$$p(k)=\frac{1}{1+A}(\frac{A}{1+A})^k\qquad k=0,1,2,\cdots$$

2.19 设 X 是[-1，1]上的均匀分布随机变量，试求 $H_C(X)$，$H_C(X^2)$。

[解]
$$p(x)=\begin{cases}\frac{1}{2} & x\in[-1,1]\\ 0 & x\notin[-1,1]\end{cases}$$

$$H_c(X)=-\int_{-\infty}^{\infty}p(x)\log p(x)dx$$
$$=[-\frac{1}{2}\log\frac{1}{2}]\cdot 2=\log^2=1(bit)$$

令 Y=X^2**，则**

$$p(y)=\begin{cases}\frac{1}{2\sqrt{y}} & y\in[0,1]\\ 0 & y\notin[0,1]\end{cases}$$

所以

$$H_C(Y)=-\int_{-\infty}^{\infty}p(y)\log p(y)dy$$
$$=-\int_0^1\frac{1}{2\sqrt{y}}\ln(\frac{1}{2\sqrt{y}})dy=\ln 2-1\ \text{nat}$$

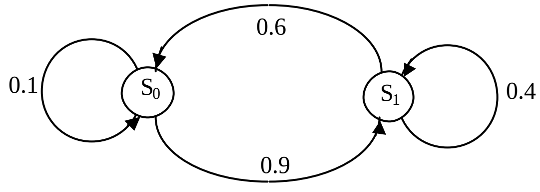
2.22 一个二状态平稳马尔可夫信源，输出为…X_0，X_1，X_2，…，信源状态转移矩阵为 $P=\begin{pmatrix}\frac{1}{10} & \frac{9}{10}\\ \frac{3}{5} & \frac{2}{5}\end{pmatrix}$**，对任何 n 求** $\frac{1}{n}H(X_1X_2\cdots X_n)$ **和** $H(X_n|X_1X_2\cdots X_{N-1})$**。**

[解] 由于信源是一阶平稳马尔可夫过程，所以

$$\frac{1}{n}H(X_1X_2\cdots X_n)=\frac{1}{n}[H(X_1)+H(X_2|X_1)+\cdots+H(X_n|X_{n-1})]$$
$$=\frac{1}{n}H(X_1)+\frac{n-1}{n}H(X_2|X_1)$$

$$H(X_n|X_{n-1}X_{n-2}\cdots X_1)=\begin{cases}H(X_1) & n=1\\ H(X_2|X_1) & n>1\end{cases}$$

由状态转移矩阵可以画出状态转移图如下：



稳态状态概率满足下面方程：

$$0.1P(S_0)+0.3P(S_1)=P(S_0)$$
$$0.9P(S_0)+0.2P(S_1)=P(S_1)$$
$$P(S_0)+P(S_1)=1$$

解出状态概率为：

$$P(S_0)=0.4, \quad P(S_1)=0.6$$

因为 $H(X\mid S_0)=0.469$ bit，$H(X\mid S_1)=0.971$ bit,

所以 $H(X_2\mid X_1)=P(S_0)H(X\mid S_0)+P(S_1)H(X\mid S_1)$

$$=0.77\ \text{bit}$$

由于无条件输出符号概率为：

$$p(0)=0.1P(S_0)+0.6P(S_1)=0.4$$
$$p(1)=0.9P(S_0)+0.4P(S_1)=0.6$$

所以 $H(X_1)=0.97$ bit

从而 $\frac{1}{n}H(X_1X_2\cdots X_n)=\frac{0.97}{n}+\frac{n-1}{n}\cdot 0.77$ bit，$n=1,2,3,\cdots$

$$H(X_n\mid X_{n-1}X_{n-2}\cdots X_1)=\begin{cases}0.97\text{bit} & n=1\\ 0.77\text{bit} & n>1\end{cases}$$

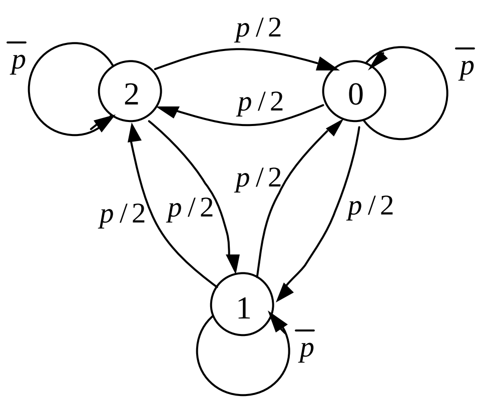
2.24 一阶马尔可夫信源的状态图如图所示，信源符号集为{0，1，2}**，并定义** $\bar{p}=1-p$

(a) **求信源平稳概率分布** $p(0), p(1), p(2)$ **；**

(b) **求此信源的熵；**

(c) **近似认为此信源为无记忆时，符号概率分布等于平稳分布，求此近似信源的熵** H（X）**，并与** H_∞ **进行比较；**

(d) **对一阶马尔可夫信源，**p **取何值时** H_∞ **取最大值；又当** p=0 **和** p=1 **时结果如何？**



习题 2.24 **图**

[解]

(a) **稳态分布满足**

$$\begin{cases}\bar{p}\cdot p(0)+\frac{p}{2}\cdot p(1)+\frac{p}{2}\cdot p(2)=p(0)\\ \frac{p}{2}\cdot p(0)+\bar{p}\cdot p(1)+\frac{p}{2}p(2)=p(1)\\ \frac{p}{2}\cdot p(0)+\frac{p}{2}\cdot p(1)+\bar{p}\cdot p(2)=p(2)\\ p(0)+p(1)+p(2)=1\end{cases}$$

由方程对称性，显然解为 $p(0)=p(1)=p(2)=\frac{1}{3}$

(b) **在状态** $S=i$ **条件下的条件熵为**

$$H(X\mid S=i)=-\bar{p}\log\bar{p}-p\log\frac{p}{2}\text{，}\quad i=0,1,2$$

则信源熵率

$$H_{\infty}=\sum_{i=0}^{2}p(S=i)\cdot H(X\mid S=i)=\bar{p}\log\bar{p}-\dot{p}\log\frac{p}{2}$$

(c) **当信源作为无记忆时，信源输出符号** i **的概率为**

$$p(x=i)=\sum_{j=0}^{3}p(S=j)\cdot p(x=i\mid S=j)$$

$$=\frac{1}{3}\text{，}\qquad\qquad i=0,1,2$$

所以近似无记忆信源熵为

$$H(X)=\log 3\geq H_{\infty}=-\bar{p}\log\bar{p}-p\log\frac{p}{2}$$

(d) **当** $p=\frac{2}{3}$**时，** H_{∞} **达到极大值为** $\log 3$ bit

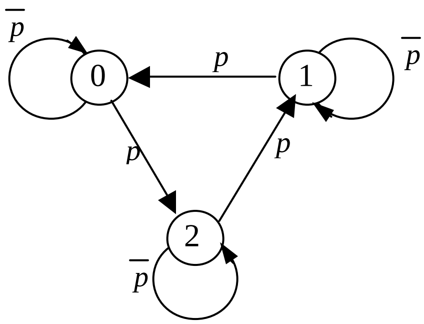
当 $p=0$ **时，** $H_{\infty}=0$ bit**；当** $p=1$ **时，** $H_{\infty}=1$ bit

2.25 一阶马尔可夫信源的状态转移图如图所示。信源 X **的符号集为**{0，1，2}

(1) **求平稳后信源的概率分布；**

(2) **求信源的熵** H_{∞}**；**

(3) **求当** $p=0$ **和** $p=1$ **时的信源熵，并说明理由。**



习题 2.25 **图**

[解]

(1) **平稳后信源状态分布** $p(0)$，$p(1)$，$p(2)$**满足**

$$\begin{cases} \bar{p}\cdot p(0)+p\cdot p(1)=p(0) \\ \bar{p}\cdot p(1)+p\cdot p(2)=p(1) \\ \bar{p}\cdot p(2)+p\cdot p(0)=p(2) \\ p(0)+p(1)+p(2)=1 \end{cases}$$

解出状态稳态分布为

$$p(0)=p(1)=p(2)=\frac{1}{3}$$

(2) **在状态** $S=i$ **条件的条件熵为**

$$H(X\mid S=i)=-\bar{p}\log\bar{p}-p\log p \qquad i=0,1,2$$

$$H_\infty=\sum_{i=0}^{2}p(i)H(X\mid S=i)=-\bar{p}\log\bar{p}-p\log p$$

(3) **当** $P=0,1$，$H_\infty=0$

这时信源无不确定性， $p=0$ **时表示信源输出常数，**$p=1$ **时表示信源周期地输出** 0，1，2。

2.5 题另解：

$$I(X;Y)=H(Y)-H(Y/X)$$

$$p(y=n)=\begin{cases}p(x=n) & ,\quad n为偶数\\ p(y/x=n)\cdot p(x=n)+p(y/x\neq n)\cdot p(x\neq n) & ，n 为奇数\end{cases}$$

$$=0.1$$

$$\therefore\quad H(Y)=-0.1\cdot\log 0.1=3.32\ bit$$

$$H(Y/X)=H(Y/X=偶数)\cdot P(X=偶数)+H(Y/X=奇数)\cdot P(X=奇数)$$

$$=0-0.5\times[0.5\times\log 0.5+4\times 0.125\times\log 0.125]$$

$$=1\ \text{bit}$$

$$\therefore\quad I(X;Y)=H(Y)-H(Y/X)=3.322-1=2.322\ \text{bit}$$

2.6　对任意概率分布的随机变量，试证明下述三角不等式成立

(a)　$H(X/Y)+H(Y/Z)\geq H(X/Z)$

$$\frac{H(X/Y)}{H(X,Y)}+\frac{H(Y/Z)}{H(Y,Z)}\geq\frac{H(X/Y)+H(Y/Z)}{H(X,Z)}$$

[提示]　$$\frac{H(X/Y)}{H(X,Y)}+\frac{H(Y/Z)}{H(Y,Z)}\geq\frac{H(X/Y)+H(Y/Z)}{H(X/Y)+H(Y/Z)+H(Z)}$$

$$\geq\frac{H(X/Z)}{H(X/Z)+H(Z)}\geq\frac{H(X/Z)}{H(X,Z)}$$

其中，第二个不等式是由于本题（a）和函数 $f(x)=\dfrac{x}{x+H(Z)}$ 的单调递增性。

[证明]

(a)　$H(X/Y)+H(Y/Z)\geq H(X/Y,Z)+H(Y/Z)=H(X,Y/Z)$,

而　$H(X,Y/Z)=H(X/Z)+H(Y/X,Z)$

$H(Y/X,Z)\geq 0$

$\therefore\quad H(X/Y)+H(Y/Z)\geq H(X/Z)$

(b)　$H(X,Y)\geq H(X/Y)+H(Y)+H(Z/Y)=H(X/Y)+H(Y/Z)+H(Z)$

$H(Y,Z)\geq H(Y/Z)+H(Z)+H(X/Y)$

又 $f(x)=\dfrac{x}{x+H(Z)}$， $f^{'}(x)=\dfrac{H(Z)}{(x+H(Z))^2}\ge 0$， $f(x)$ 单调递增

$$\therefore \quad \frac{H(X/Y)}{H(X,Y)}+\frac{H(Y/Z)}{H(Y,Z)}\ge\frac{H(X/Y)+H(Y/Z)}{H(X/Y)+H(Y/Z)+H(Z)}$$

$$\ge\frac{H(X/Z)}{H(X/Z)+H(Z)}\ge\frac{H(X/Z)}{H(Z)}$$

2.10　令 X 是离散随即变量, $Y=g(X)$ 是 X 的函数，求证 $H(X)\ge H(Y)$。

[证明] $\because$　Y 是 X 的函数

$\therefore$　$P(Y/X)=1$， $H(Y/X)=0$

又 $\because$　$H(X,Y)=H(X)+H(Y/X)=H(Y)+H(X/Y)$

$\therefore$　$H(Y)+H(X/Y)=H(X)$

$\therefore$　$H(X)\ge H(Y)$

2.11　试证明，如 $H(Y/X)=0$，则 Y 是 X 的函数，也就是对一切使 $p(x)>0$ 的 x，仅有一个可能的 y 取值，具有 $p(x,y)>0$。

[证明] 若 $H(Y/X)=0$，则对任意使 $p(x)>0$ 的 x_i 有：

$H(Y/x_i)=0$

由熵的非负性可知：

$H(Y/x_i)=-\sum_i p(y/x_i)\log(p(y/x_i))=0$

$\therefore$　$p(y/x_i)=0$ 或 $p(y/x_i)=1$

也就是任何使 $p(x)>0$ 的 x_i，都有确定的 y 取值。

2.12　令 X_1 和 X_2 为分别定义在字符表 $\chi_1=\{1,2,\cdots,m\}$ 和 $\chi_2=\{m+1,\cdots,n\}$ 上的 2 个离散随机变量，它们的分布函数为 $p_1(\bullet)$ 和 $p_2(\bullet)$。现构造随机变量

$$X=\begin{cases}X_1 & \text{以概率 } a\\ X_2 & \text{以概率 } 1-a\end{cases}$$

(a) 试用 $H(X_1)$，$H(X_2)$ 和 a 来表示 X 的熵 $H(X)$。

(b) 在 a 上求 $H(X)$ 的极大值，证明 $2^{H(X)}\le 2^{H(X_1)}+2^{H(X_2)}$。

[解]

(a) 由熵的可加性可得：

$$H(X)=aH(X_1)+(1-a)H(X_2)+H(a)$$

$$=aH(X_1)+(1-a)H(X_2)-a\log a-(1-a)\log(1-a)$$

(b) $\dfrac{dH(X)}{da}=H(X_1)-H(X_2)-\log a+\log(1-a)=0$

解得：$a=\dfrac{2^{H(X_1)-H(X_2)}}{1+2^{H(X_1)-H(X_2)}}=\dfrac{2^{H(X_1)}}{2^{H(X_1)}+2^{H(X_2)}}\qquad a\in(0,1)$

而 $\dfrac{d^2H(X)}{da^2}=\dfrac{-1}{(1-a)\ln 2}-\dfrac{1}{a\ln 2}<0$

∴ $H(X)$ 在 a 处有极大值（或者由 $H(X)$ 的凸函数性质直接得出）

$$H(X)\Big|_{\max}=\frac{2^{H(X_1)}}{2^{H(X_1)}+2^{H(X_2)}}\left[H(X_1)-\log\frac{2^{H(X_1)}}{2^{H(X_1)}+2^{H(X_2)}}\right]$$

$$+\frac{2^{H(X_2)}}{2^{H(X_1)}+2^{H(X_2)}}\left[H(X_2)-\log\frac{2^{H(X_2)}}{2^{H(X_1)}+2^{H(X_2)}}\right]$$

$$=\log[2^{H(X_1)}+2^{H(X_2)}]$$

∴ $2^{H(X)}\le 2^{H(X_1)}+2^{H(X_2)}$

2.20 令 X 是取值 ±1 的二元随机变量，概率分布为 $p(x=1)=p(x=-1)=0.5$，令 Y 是连续随机变量，已知条件概率密度为

$$p(y/x)=\begin{cases}1/4, & -2<y-x\le 2\\ 0 & \text{其他}\end{cases}$$

试求

(a) Y 的概率密度。

(b) $I(X;Y)$。

(C) 若对 Y 做硬判决 $V=\begin{cases}1, & y>1\\ 0 & -1<y\le 1\\ -1 & y\le -1\end{cases}$

求 $I(V;X)$，并对结果加以解释。

[解]

(a) $p(y/x=1)=\begin{cases}1/4 & -1<y\le 3\\ 0 & \text{其他}\end{cases}$

$$p(y/x=-1)=\begin{cases}1/4 & -3<y\le 1\\ 0 & \text{其他}\end{cases}$$

而 $p(y)=p(y/x=1)\cdot p(x=1)+p(y/x=-1)\cdot p(x=-1)$

$$\therefore \quad p(y)=\begin{cases}1/8 & -3<y\le -1,\ 1<y\le 3\\ 1/4 & -1<y\le 1\\ 0 & \text{其他}\end{cases}$$

(b) $I(X;Y)=\sum_x\int_y p(x,y)\cdot\log\dfrac{P(y/x)}{p(y)}dy$

$$=\sum_x\int_y p(y/x)\cdot p(x)\log\frac{P(y/x)}{p(y)}dy$$

$$=\frac{1}{2}\int_{-1}^{3}\frac{1}{4}\log\frac{1}{2}dy+\frac{1}{2}\int_{-3}^{-1}\frac{1}{4}\log\frac{1}{2}dy$$

$$+\frac{1}{2}\int_{-1}^{-1}\frac{1}{4}\log\frac{1/4}{1/4}dy+\frac{1}{2}\int_{-1}^{-1}\frac{1}{4}\log\frac{1/4}{1/4}dy$$

$$=\frac{1}{2}$$

(c) $p(v=1)=1/4$， $p(v=0)=1/2$， $p(v=-1)=1/4$

$$p(v/x=1)=\begin{cases}1/2 & \text{v=1}\\ 1/2 & \text{v=0}\\ 0 & \text{v}=-1\end{cases}$$

$$p(v/x=-1)=\begin{cases} 0 & \text{v=1} \\ 1/2 & \text{v=0} \\ 1/2 & \text{v}=-1 \end{cases}$$

$$I(V;X)=\sum_{x}\sum_{v} p(x,v)\cdot\log\frac{P(v/x)}{p(v)}$$

$$=\sum_{x}\sum_{v} p(v/x)\cdot p(x)\cdot\log\frac{P(v/x)}{p(v)}$$

$$=\frac{1}{2}[\frac{1}{2}\cdot\log\frac{1/2}{1/4}+\frac{1}{2}\cdot\log\frac{1/2}{1/2}+\frac{1}{2}\cdot\log\frac{1/2}{1/2}+\frac{1}{2}\cdot\log\frac{1/2}{1/4}]$$

$$=\frac{1}{2}$$

$\therefore\quad I(V;X)=I(X;Y)$ ，即该硬判决没有损失信号。