第2章 习题

2.2. 函数的熵。设X是取有限个值的随机变量，如果

(a) $Y=2^X$

(b) $Y=\cos X$

$H(X)$和$H(Y)$的不等关系是什么？

(a)：$H(X)=H(Y)$

(b)：$H(X)\geqslant H(Y)$

2.3 最小熵。求$H(p_1,p_2,\cdots p_n)=H(\vec{p})$的最小值，其中$\vec{p}$的取值域为$n$维概率向量集合。

请找出所有达到这个最小值时的$\vec{p}$.

$\vec{p}$是I_n的列向量（I_n是$n\times n$单位矩阵）

2.4 随机变量的函数的熵。设X为离散型随机变量，证明X的函数的熵必小于等于X的熵

$$H(X,g(X))=H(X)+H(g(X)|X)$$

$$=H(X)$$

$\because \log \underline{p(g(x)|x)}=0$　$\nearrow 1$

$$H(X,g(X))=H(g(X))+\underbrace{H(X|g(X))}_{\geqslant 0}$$

$$\geqslant H(g(X))$$

$$\therefore H(X)\geqslant H(g(X))$$

2.5 零条件熵。证明：若$H(Y|X)=0$，则Y是X的函数。

（即对于满足$p(x)>0$的任意x，仅存一个可能取值y，使$p(x,y)>0$）

反证法：若$\exists\, y_1,y_2$ 且$y_1\neq y_2$ 使得对$\forall x$，$p(x,y)>0$

则 $p(x)>0$，$p(x,y_1)>0$，$p(x,y_2)>0$

$$H(Y|X)=\sum_{x\in X}\cancel{P(X=}\ p(x)\cdot H(Y|X=x)$$

$$\geqslant p(x)\cdot H(Y|X=x)$$

$$=p(x)\Big(-\sum_{y\in Y}p(y|x)\log p(y|x)\Big)$$

$$\geqslant p(x)\cdot\big(-p\overset{(y_1|x)}{\cancel{(x,y_1)}}\log p\overset{(y_1|x)}{\cancel{(x,y_1)}}-p\overset{(y_2|x)}{\cancel{(x,y_2)}}\log p\overset{y_2|x}{\cancel{(x,y_2)}}\big)$$

$$>0$$

由题意知$H(Y|X)=0$ $\therefore$ 假设错误，原结论得证

2.10 不相交组合的熵。设离散型随机变量 X_1 和 X_2 的概率密度函数分别为 $p_1(\cdot)$ 和 $p_2(\cdot)$，字母表分别为 $\mathcal{X}_1=\{1,2,\cdots,m\}$，$\mathcal{X}_2=\{m+1,\cdots,n\}$。设

$$X=\begin{cases} X_1, & 概率为 \alpha \\ X_2, & 概率为 1-\alpha \end{cases}$$

(a) 试求 $H(X)$ 关于 $H(X_1)$，$H(X_2)$ 和 α 的表达式。

(b) α 为何值时，$H(X)$ 最大.

2.13 不等式. 证明对任意的 $x>0$，$\ln x \geq 1-\frac{1}{x}$。

记 $f(x)=\ln x-1+\frac{1}{x}$，$x>0$

$f'(x)=\frac{1}{x}-\frac{1}{x^2}=\frac{x-1}{x^2}$

$f'(x)>0 \Rightarrow x>1$

$f'(x)<0 \Rightarrow 0<x<1$

$\therefore f(x)$ 在 $(0,1]\downarrow$ $[1,+\infty)\uparrow$

$\therefore f(x)\geq f(1)=0$

$\therefore$ 当 $x>0$ 时，$\ln x\geq 1-\frac{1}{x}$

2.12. 联合熵的例子. 设 $p(x,y)$ 由右表给出，试计算

(a) $H(X)$，$H(Y)$。

(b) $H(X|Y)$，$H(Y|X)$

(c) $H(X,Y)$

(d) $H(Y)-H(Y|X)$

(e) $I(X;Y)$

X \ Y	0	1	$p(x)$
0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$
1	0	$\frac{1}{3}$	$\frac{1}{3}$
$p(y)$	$\frac{1}{3}$	$\frac{2}{3}$	

(a) $H(X)=-\frac{2}{3}\log\frac{2}{3}-\frac{1}{3}\log\frac{1}{3}$

$H(Y)=H(X)$

(b) $H(X|Y)=\sum_{i=1}^{2}p(y_i)\cancel{\sum_{j=1}^{2}\frac{p(x_j|y_i)}{p(x_j|y_i)}}$ $H(X|Y=y_i)$

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

$=\frac{2}{3}$

$H(Y|X)=\sum_{i=1}^{2}p(x_i)\ H(Y|X=x_i)$

$=\frac{2}{3}\times 1+\frac{1}{3}\times 0$

$=\frac{2}{3}$

(c) $H(X,Y)=-\frac{1}{3}\log\frac{1}{3}$ [illegible] $\log 3$

(d) $H(Y)-H($[illegible]$)=-\frac{2}{3}\log\frac{2}{3}-\frac{1}{3}\log\frac{1}{3}-\frac{2}{3}$

(e) $I(X;Y)=$ [illegible] $H(Y|X)$

2.14 和的熵. 设随机变量X,Y的取值分别为$x_1,x_2,\cdots,x_r$和$y_1,y_2,\cdots,y_s$，设$Z=X+Y$。 3

(a) 证明$H(Z|X)=H(Y|X)$，并讨论如果X,Y独立，则$H(Y)\le H(Z)$及$H(X)\le H(Z)$，由此说明独立随机变量的和增加不确定度。

(b) 给出一个(必须是相关)随机变量的例子，使$H(X)>H(Z)$且$H(Y)>H(Z)$

(c) 在什么条件下，$H(Z)=H(X)+H(Y)$？

(a) $H(Z|X)=\sum_{i=1}^{r}p(x_i)H(Z|X=x_i)$

$=-\sum_{i=1}^{r}p(x_i)\sum_{j=1}^{s}p(z_j|x_i)\log p(z_j|x_i)$

$\because Z=X+Y$

$=-\sum_{i=1}^{r}p(x_i)\sum_{j=1}^{s}p(y_j|x_i)\log p(y_j|x_i)$

$=H(Y|X)$

如果X和Y是独立的，则$H(Y|X)=H(Y)$

$I(X;Z)=H(Z)-H(Z|X)\ge 0$

$\therefore H(Z)\ge H(Z|X)$

$=H(Y|X)$

$=H(Y)$ $\because$若X,Y独立

~~$I(X,Y;Z)=H(Z)-H(Z|Y)\ge 0$~~

~~$I(X;Z)=H(X)-$~~

同理$H(Z)\ge H(X)$

(b) $X=\begin{cases}-1, & p=\frac{1}{2}\\ 1, & p=\frac{1}{2}\end{cases}$ $H(X)=1$

$Y=-X$ $H(Y)=1$

$Z=0$ $p=1$ $H(Z)=0$

(c) $\because Z$是X和Y的函数

$\therefore H(Z)\le H(X,Y)$

又$H(X,Y)=H(X)+H(Y|X)\le H(X)+H(Y)$

即$H(Z)\le H(X,Y)\le H(X)+H(Y)$

要使等号成立，则

~~Z~~ X,Y是Z的函数 且 $H(Y|X)=H(Y)$

$\therefore$ X与Y独立

4

2.15 数据处理，设 $X_1\to X_2\to X_3\to\cdots\to X_n$ 依序构成马尔可夫链，即设

$$p(x_1,x_2,\cdots,x_n)=p(x_1)\cdot p(x_2|x_1)\cdots p(x_n|x_{n-1})$$

试将 $I(X_1;X_2,X_3,\cdots,X_n)$ 简化到最简单形式.

① $p(x_1,x_2,\cdots,x_n)=p(x_1)\cdot p(x_2|x_1)\cdot p(x_3|x_1x_2)\cdots p(x_n|x_1x_2\cdots x_{n-1})$

由马尔可夫链得

$$=p(x_1)\cdot p(x_2|x_1)\cdot p(x_3|x_2)\cdots p(x_n|x_{n-1})$$

$I(X_1;X_2,X_3,\cdots,X_n)$

$=H(X_1)+H(X_2,X_3,\cdots,X_n)-H(X_1,X_2,\cdots,X_n)$

$=H(X_1)+(H(X_2)+H(X_3|X_2)+H(X_4|X_3X_2)+\cdots+H(X_n|X_{n-1}\cdots X_1))-(H(X_1)+H(X_2|X_1)+H(X_3|X_2X_1)+\cdots+H(X_n|X_{n-1}\cdots X_1))$

$=H(X_1)+(H(X_2)+H(X_3|X_2)+H(X_4|X_3)+\cdots+H(X_n|X_{n-1}))-(H(X_1)+H(X_2|X_1)+H(X_3|X_2)+\cdots+H(X_n|X_{n-1}))$

$=H(X_2)-H(X_2|X_1)$

$=I(X_1;X_2)$

另一种方法：

$$I(X_1;X_2,X_3)=\sum_{x_1x_2x_3}p_{x_1x_2x_3}\log\frac{p_{x_1x_2x_3}}{p_{x_1}p_{x_2x_3}}$$

$$=\sum_{x_1x_2x_3}p_{x_1|x_2x_3}p_{x_2x_3}\log\frac{p_{x_1|x_2x_3}}{p_{x_1}}$$

$$=\sum_{x_2x_3}p_{x_2x_3}\sum_{x_1}p_{x_1|x_2}\log\frac{p_{x_1|x_2}}{p_{x_1}}$$

$$=\sum_{x_1x_2}p_{x_2}p_{x_1|x_2}\log\frac{p_{x_1|x_2}}{p_{x_1}}$$

$$=\sum_{x_1x_2}p_{x_1x_2}\log\frac{p_{x_1x_2}}{p_{x_1}p_{x_2}}$$

$$=I(X_1;X_2)$$

2.27. 熵的组合法则。设$\vec{p}=(p_1,p_2,\cdots,p_m)$为$m$个元素上的概率分布(即$p_i\geq 0$，且$\sum_{i=1}^{m}p_i=1$)。
定义分布$\vec{q}$与$\vec{p}$在集合$\{1,2,\cdots,m-2\}$上是相同的，$\vec{q}$中最后一个元素的概率为$\vec{p}$中最后两个元素的概率之和。证明：

$$H(\vec{p})=H(\vec{q})+(p_{m-1}+p_m)\cdot H\left(\frac{p_{m-1}}{p_{m-1}+p_m},\ \frac{p_m}{p_{m-1}+p_m}\right)$$

$$H(\vec{p})-H(\vec{q})=-p_{m-1}\log p_{m-1}-p_m\log p_m-(-(p_{m-1}+p_m)\log(p_{m-1}+p_m))$$

$$=-p_{m-1}\log\frac{p_{m-1}}{p_m+p_{m-1}}-p_m\log\frac{p_m}{p_{m-1}+p_m}$$

$$=(p_{m-1}+p_m)\,H\left(\frac{p_{m-1}}{p_{m-1}+p_m},\frac{p_m}{p_{m-1}+p_m}\right)$$

2.28 混合使熵增加。证明概率分布$\underbrace{(p_1,p_2,\cdots,p_j,\cdots,p_m)}_{\to H_1}$的熵小于概率分布$\underbrace{(p_1,\cdots,\frac{p_i+p_j}{2},\cdots,\frac{p_i+p_j}{2},\cdots,p_m)}_{\to H_2}$的熵。进一步证明更一般的结论：使概率分布更均匀的变换都使熵增加。

$$H_2-H_1=-\frac{p_i+p_j}{2}\times 2\log\frac{p_i+p_j}{2}+p_i\log p_i+p_j\log p_j$$

$$=p_j\log\frac{2p_j}{p_i+p_j}+p_i\log\frac{2p_i}{p_i+p_j}$$

设$p_i+p_j=m$ $0<m<1$
$p_i=x$ 则 $p_j=m-x$

则 $H_2-H_1=(m-x)\log\frac{2(m-x)}{m}+x\log\frac{2x}{m}$

$(H_2-H_1)'=\log\frac{x}{m-x}$

当$x<\frac{m}{2}$时，$(H_2-H_1)'<0$ ↓

$x>\frac{m}{2}$时，$(H_2-H_1)'>0$ ↑

∴当$x=\frac{m}{2}$时 $(H_2-H_1)_{min}=0$

∴ $H_2-H_1\geq 0$

2.29 不等式。设X、Y和Z为联合随机变量，~~对固定的值A≥0~~，证明下列不等式，并给出等号成立条件

条件：Y是X和Z的函数

(a) $H(X,Y|Z)\geq H(X|Z)$ (a) $H(X,Y|Z)=H(X|Z)+\underbrace{H(Y|X,Z)}_{\geq 0}\geq H(X|Z)$

(b) $I(X,Y;Z)\geq I(X;Z)$ (b) $I(X,Y;Z)-I(X;Z)=H(X,Y)+H(Z)-$[illegible]$-$

(c) $H(X,Y,Z)-H(X,Y)\leq H(X,Z)-H(X)$ $(H(X)+H(Z)-H(X,Z))$

(d) ~~I(Z;~~ $I(X;Z|Y)\geq I(Z;Y|X)-I(Z;Y)+I(X;Z)$ $=$[illegible]$-H(X)+H(X,Z)$

$-H(X,Y,Z)$

条件：给定X时，Y和Z条件独立 $I(Y;Z|X)$ $=H(Y|$[illegible]$)+$[illegible]$(Y|X,Z)\geq 0$

(d) $I(X;Z|Y) \geq I(Z;Y|X) - I(Z;Y) + I(X;Z)$

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

$$I(X;Z|Y) - I(Z;Y|X)$$
$$= H(Y,Z) - H(Y) + H(X) - H(X,Z)$$
$$I(X;Z) - I(Z;Y)$$
$$= H(X) - H(X,Z) - H(Y) + H(Y,Z)$$
$$\therefore I(X;Z|Y) - I(Z;Y|X) = I(X;Z) - I(Z;Y)$$
即 $I(X;Z|Y) = I(Z;Y|X) - I(Z;Y) + I(X;Z)$

(c) $H(X,Y,Z) - H(X,Y) \leq H(X,Z) - H(X)$ ☆☆☆

$$H(X,Y,Z) - H(X,Y) = H(Z|X,Y)$$
$$= H(Z|X) - I(Y;Z|X)$$
$$\leq H(Z|X)$$
$$= H(X,Z) - H(X)$$

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

条件：给定X下，Y和Z条件独立。

2.30 最大熵。设X是取非负整数值的随机变量，对固定的值$A>0$，试求在约束条件 6

$EX=\sum_{n=0}^{\infty} np(n)=A$ 下使得熵$H(X)$达到最大时的概率密度函数$p(x)$，

并计算出$H(X)$的最大值.

2.42 不等式。"$\geq$"，"$\leq$"，"$=$"

(a) $H(5X)$与$H(X)$ $H(5X)=H(X)$

(b) $I(g(X);Y)$与$I(X;Y)$

(c) $H(X_0|X_{-1})$与$H(X_0|X_{-1},X_1)$ $H(X_0|X_{-1})\geq H(X_0|X_{-1},X_1)$

(d) $\dfrac{H(X,Y)}{H(X)+H(Y)}$与1。 $\dfrac{H(X,Y)}{H(X)+H(Y)}\leq 1$