2022 春信息论 B 第一次习题课补充题目

考察范围：第二章 ~ 第五章　　课程助教：高源

1. 多选题。设 X,Y,Z 为离散随机变量，U,V 为连续随机变量，则以下等式或不等式正确的是 (AB)

$$
\begin{aligned}
&(A)H(g(X)) \leqslant H(X)\\
&(B)\text{若}X,Y\text{独立，则有}H(XY)=H(X)+H(Y)\\
&(C)h(U) \geqslant I(U;V)\\
&(D)I(X;Y|Z) \leqslant I(X;Y)
\end{aligned}
$$

解析：

(A) 正确. 为了比较二者的关系，我们首先要在二者之间建立联系，即考察二者的联合熵 $H(X,g(X))$. 我们对联合熵 $H(X,g(X))$ 尝试**两种不同方式的分解**，进而可以建立 $H(X)$ 和 $H(g(X))$ 的联系。

第一种分解

$$
\begin{aligned}
H(X,g(X)) &\overset{(a)}{=} H(X)+H(g(X)\mid X)\\
&\overset{(b)}{=} H(X)
\end{aligned}
$$

第二种分解

$$
\begin{aligned}
H(X,g(X)) &\overset{(c)}{=} H(g(X))+H(X\mid g(X))\\
&\geqslant H(g(X))
\end{aligned}
$$

其中

(a) 由熵的链式法则可知.

(b) 由于给定了 $X=x$ 后，$g(X)=g(x)$ 为定值，所以条件熵为 0.

(c) 由熵的链式法则可知.

(d) 由于条件熵非负. 等号成立条件为：$g(x)$ 为可逆映射.

由上述讨论可知

$$H(g(X)) \leqslant H(X)$$

可得结论正确。

地址 (Address): 安徽省合肥市金寨路 96 号　　邮编 (Postcode):230026
电子邮箱 (E-mail):wazs98@mail.ustc.edu.cn　　个人主页 (Homepage):home.ustc.edu.cn/ wazs98

(B) 正确. 设两个统计独立的信源 X 和 Y 的信源空间分别为

$$[X \cdot P]:\left\{\begin{array}{rcccc} X: & a_1 & a_2 & \cdots & a_r \\ P(X): & p_1 & p_2 & \cdots & p_r \end{array}\right.$$

其中 $0 \leqslant p_i \leqslant 1(i=1,2,\ldots,r)$，$\sum_{i=1}^{r} p_i = 1$.

$$[Y \cdot P]:\left\{\begin{array}{rcccc} Y: & b_1 & b_2 & \cdots & b_s \\ P(Y): & q_1 & q_2 & \cdots & q_s \end{array}\right.$$

其中 $0 \leqslant q_j \leqslant 1(j=1,2,\ldots,s)$，$\sum_{i=1}^{s} q_j = 1$。

若信源 X 和 Y 同时发出各自的符号，则构成联合信源 (XY)，其符号表为

$$(XY):\{(a_i b_j)\,(i=1,2,\ldots,r;j=1,2,\ldots,s)\}$$

概率分布为

$$p\left(a_i b_j\right)=p\left(a_i\right) p\left(b_j\right)=p_i q_j \quad (i=1,2,\ldots,r;j=1,2,\ldots,s)$$

满足

$$\sum_{i=1}^{r} \sum_{j=1}^{s} p\left(a_i b_j\right)=\sum_{i=1}^{r} \sum_{j=1}^{s} p\left(a_i\right) p\left(b_j\right)=\sum_{i=1}^{r} \sum_{j=1}^{s} p_i q_j=1$$

所以有

$$\begin{aligned}
H(XY) &= -\sum_{i=1}^{r}\sum_{j=1}^{s} p\left(a_i b_j\right)\log p\left(a_i b_j\right) \\
&= -\sum_{i=1}^{r}\sum_{j=1}^{s} p_i q_j \log p_i q_j \\
&= -\sum_{i=1}^{r}\sum_{j=1}^{s} p_i q_j \log p_i - \sum_{i=1}^{r}\sum_{j=1}^{s} p_i q_j \log q_j \\
&= \sum_{j=1}^{s} q_j\left\{-\sum_{i=1}^{r} p_i \log p_i\right\} + \sum_{i=1}^{r} p_i\left\{-\sum_{j=1}^{s} q_j \log q_j\right\} \\
&= \sum_{j=1}^{s} q_j H\left(p_1, p_2, \ldots, p_r\right) + \sum_{i=1}^{r} p_i H\left(q_1, q_2, \ldots, q_s\right) \\
&= \sum_{j=1}^{s} q_j H(X) + \sum_{i=1}^{r} p_i H(Y) \\
&= H(X) + H(Y)
\end{aligned}$$

所以命题成立. 实际上，这个命题叫做熵函数可加性定理. 熵函数的可加性表明, 由两个相互统计独立的离散信源 X 和 Y 构成的联合信源 (XY) 的熵，等于信源 X 和 Y 各自熵之和. 联合信源 (XY) 每发一个消息所能提供的平均信息量, 等于信源 X 和 Y 各自发一个符号提供的平均信息量之和 (当 X 和 Y 不统计独立, 相互间存在统计依赖关系时, 熵函数的可加性同样成立). 熵函数的可加性的数学意义是：**两个熵函数之和仍然是一个熵函数**.

(C) 错误. 反例：U 为 $[0,\frac{1}{2}]$ 上的均匀分布，$h(U)=\log\frac{1}{2}=-1$ 比特，对于 V 服从任意分布，都有 $I(U;V)\geqslant 0$. 所以命题错误.

(D) 错误." 条件使得互信息减少" 这一结论是有成立条件的,要求 $X\to Y\to Z$ 构成马尔可夫链. 若对于一般的情形，存在 X,Y,Z 使得 $I(X;Y|Z)\geqslant I(X;Y)$.

考虑 X,Y 独立情况下，$I(X;Y)=0$. 记 $Z=X+Y$，有

$$I(X;Y\mid Z)=H(X\mid Z)-H(X\mid Y,Z)=H(X\mid Z)\geqslant 0$$

若 Y 不是常量，则可以进一步得 $I(X;Y\mid Z)>I(X;Y)=0$.

2. 判断题 (若判断为对，简要说明或证明；若判断为错，简要说明或举出反例)

假设随机变量 X，Y 和 Z,W 构成如下的马尔可夫链

$$X \to Y \to (Z,W),\ 即p(x,y,z,w) = p(x)p(y \mid x)p(z,w \mid y)$$

则有

$$I(X;Z) + I(X;W) \leqslant I(X;Y) + I(Z;W)$$

证明：由数据处理不等式知

$$I(X;Y) \geqslant I(X;(Z,W))$$

于是有

$$\begin{aligned}
&I(X;Y) + I(Z;W) - I(X;Z) - I(X;W) \\
&\qquad \geqslant I(X;Z,W) + I(Z;W) - I(X;Z) - I(X;W) \\
&\qquad = H(Z,W) + H(X) - H(X,W,Z) + H(W) + H(Z) - H(W,Z) \\
&\qquad\quad - H(X) - H(Z) + H(X,Z) - H(X) - H(W) + H(X,W) \\
&\qquad = -H(X,W,Z) + H(X,Z) + H(X,W) - H(X) \\
&\qquad = H(W \mid X) - H(W \mid X,Z) \\
&\qquad = I(W;Z \mid X) \geqslant 0
\end{aligned}$$

3. 每帧电视图像可以认为是由 3×10^5 个像素组成, 所以像素均是独立变化，且每一像素又取 128 个不同的亮度电平, 并设亮度电平等概率出现。问每帧图像含有多少信息量? 若现有一广播员在约 10000 个汉字的字汇中选 1000 个来口述此电视图像, 试问广播员描述此图像所广播的信息量是多少（假设汉字是等概率分布，并且彼此无依赖）? 若要恰当地描述此图像, 广播员在口述中至少需用多少汉字?

解析：每个像素的电平亮度形成了一个概率空间

$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \cdots & a_{128} \\ \frac{1}{128} & \frac{1}{128} & \cdots & \frac{1}{128} \end{bmatrix}$$

平均每个像素携带的信息量为：

$$H(X) = \log 128 = 7\ 比特/像素$$

地址 (Address): 安徽省合肥市金寨路 96 号
电子邮箱 (E-mail):wazs98@mail.ustc.edu.cn
邮编 (Postcode):230026
个人主页 (Homepage):home.ustc.edu.cn/ wazs98

每帧图像由 3×10^5 个像素组成, 且像素间是独立的, 因此每帧图像含有的信息量为:

$$H\left(X^N\right) = NH(X) = 2.1 \times 10^6 \text{ 比特/帧}$$

如果用汉字来描述此图像, 平均每个汉字携带的信息量为 $H(Y) = \log 10000 = 13.29$ 比特 /汉字, 选择 1000 字来描述, 携带的信息量为

$$H\left(Y^N\right) = NH(Y) = 1.329 \times 10^4 \text{ 比特}$$

如果要恰当的描述此图像, 即信息不丢失, 在上述假设不变的前提下, 需要的汉字个数

$$\frac{H\left(X^N\right)}{H(Y)} = \frac{2.1106}{13.29} \approx 1.58 \times 10^5 \text{ 字}$$

4. 设有一批电阻, 按阻值分 70% 是 2kΩ, 30% 是 5kΩ; 按功耗分 64% 是 1/8 W, 其余是 1/4 W 。现已知 2kΩ 阻值的电阻中 80% 是 1/8 W 。问通过测量阻值可以平均得到的关于瓦数的信息量是多少?

解析：根据已知条件, 设电阻的阻值为 X, 电阻的功耗为 Y, 则两随机变量的概率分布为:

$$\begin{bmatrix} X \\ P \end{bmatrix} = \begin{bmatrix} x_1 = 2k\Omega & x_2 = 5k\Omega \\ 0.7 & 0.3 \end{bmatrix}, \begin{bmatrix} Y \\ P \end{bmatrix} = \begin{bmatrix} y_1 = 1/8W & y_2 = 1/4W \\ 0.64 & 0.36 \end{bmatrix}$$

给定条件为 $P\left(y_1 \mid x_1\right) = 0.8, P\left(y_2 \mid x_1\right) = 0.2$，可得

$$\begin{aligned} 0.64 = P\left(y_1\right) = P\left(x_1\right) P\left(y_1 \mid x_1\right) + P\left(x_2\right) P\left(y_1 \mid x_2\right) = 0.7 \times 0.8 + 0.3 \times P\left(y_1 \mid x_2\right) \\ 0.36 = P\left(y_2\right) = P\left(x_1\right) P\left(y_2 \mid x_1\right) + P\left(x_2\right) P\left(y_2 \mid x_2\right) = 0.7 \times 0.2 + 0.3 \times P\left(y_2 \mid x_2\right) \end{aligned}$$

解得

$$P\left(y_1 \mid x_2\right) = \frac{4}{15}, \quad P\left(y_2 \mid x_2\right) = \frac{11}{15}$$

进而可得

$$H(Y \mid X) = -0.7 \times (0.8 \log 0.8 + 0.2 \log 0.2) - 0.3 \times \left(\frac{4}{15} \log \frac{4}{15} + \frac{11}{15} \log \frac{11}{15}\right) = 0.7567 \text{比特}$$

所以

$$I(X;Y) = H(Y) - H(Y \mid X) = 0.186 \text{ 比特}$$

5. 随机变量 X, Y 的取值分别为 $x_1, x_2, \ldots, x_m$ 和 $y_1, y_2, \ldots, y_n$。设 $Z = X + Y$。试证明 $H(Z) \leqslant H(X, Y)$。

   解析：构造信息量的展开式

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

   由于

$$H(Z \mid X, Y) = 0, H(X, Y \mid Z) \geqslant 0$$

   可得

$$H(Z) \leqslant H(X + Y)$$

6. 设离散无记忆信源

$$\begin{bmatrix} X \\ P(x) \end{bmatrix} = \begin{bmatrix} a_1 = 0 & a_2 = 1 & a_3 = 2 & a_4 = 3 \\ 3/8 & 1/4 & 1/4 & 1/8 \end{bmatrix}$$

   其发出的消息为 (202120130213001203210110321010021032011223210)，求在此消息中平均每个符号携带的信息量是多少？

   解析：由题意知，信源是无记忆的。因此，发出的各消息之间是互相独立的，此时发出的消息的自信息即为各消息的自信息之和。根据已知条件，发出各消息所包含的信息量分别为

$$I\left(a_0 = 0\right) = \log \frac{8}{3} = 1.415\text{比特}$$

$$I\left(a_1 = 1\right) = \log 4 = 2\text{比特}$$

$$I\left(a_2 = 2\right) = \log 4 = 2\text{比特}$$

$$I\left(a_3 = 3\right) = \log 8 = 3\text{比特}$$

   在发出的消息中，共有 14 个“0”符号、13 个“1”符号、12 个“2 ”符号和 6 个“3 ”符号，则得到消息的自信息为

$$I = 14 \times 1.415 + 13 \times 2 + 12 \times 2 + 6 \times 3 \approx 87.81 \text{ 比特}$$

地址 (Address): 安徽省合肥市金寨路 96 号
邮编 (Postcode):230026
电子邮箱 (E-mail):wazs98@mail.ustc.edu.cn
个人主页 (Homepage):home.ustc.edu.cn/ wazs98

所以平均每个符号携带的信息量为

$$I = \frac{87.81}{45} = 1.95 \text{ 比特/符号}$$

地址 (Address): 安徽省合肥市金寨路 96 号
电子邮箱 (E-mail):wazs98@mail.ustc.edu.cn
邮编 (Postcode):230026
个人主页 (Homepage):home.ustc.edu.cn/ wazs98