信息论作业 2

史泽宇

2020 年 3 月 4 日

题目 1 这里令 X_1 表示第一颗骰子的结果，X_2 表示第二颗骰子的结果，X_3 表示第三颗骰子的结果，所以 $X = X_1, Y = X_1 + X_2, Z = X_1 + X_2 + X_3$。

$$\begin{aligned}
H(X) &= H(X_1) = H(X_2) = H(X_3) && (1)\\
&= \sum_{1}^{6} \frac{1}{6} \log_2 6 && (2)\\
&= \log_2 6 bit && (3)\\
&\approx 2.5850 bit && (4)\\
& && (5)\\
H(Y) &= H(X_1 + X_2) = H(X_2 + X_3) = H(X_1 + X_3) && (6)\\
&= \sum_{p}^{\Omega} p \log_2 \frac{1}{p} && (7)\\
&\approx 3.2744 bit && (8)\\
\Omega &= \{\frac{1}{36}, \frac{2}{36}, \frac{3}{36}, \frac{4}{36}, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \frac{4}{36}, \frac{3}{36}, \frac{2}{36}, \frac{1}{36}\} && (9)\\
& && (10)\\
H(Z) &= H(X_1 + X_2 + X_3) && (11)\\
&= \sum_{p}^{\Omega} p \log_2 \frac{1}{p} && (12)\\
&\approx 3.5993 bit && (13)\\
\Omega &= \{\frac{1}{216}, \frac{3}{216}, \frac{6}{216}, \frac{10}{216}, \frac{15}{216}, \frac{21}{216}, \frac{25}{216}, \frac{27}{216}, \frac{27}{216}, \frac{25}{216}, \frac{21}{216}, \frac{15}{216}, \frac{10}{216}, \frac{6}{216}, \frac{3}{216}, \frac{1}{216}\} && (14)
\end{aligned}$$

1.

$$\begin{aligned}
I(Y;Z) &= H(Z) - H(Z|Y) && (15)\\
&= H(Z) - H(X) && (16)\\
&\approx 1.0143 bit && (17)
\end{aligned}$$

2.

$$\begin{aligned} I(X;Z) &= H(Z) - H(Z|X) && (18) \\ &= H(Z) - H(Y) && (19) \\ &\approx 0.3249bit && (20) \end{aligned}$$

3.

$$\begin{aligned} I(X,Y;Z) &= I(X;Z) + I(Y;Z|X) && (21) \\ &= H(Z) - H(Y) + H(Z|X) - H(Z|X,Y) && (22) \\ &= H(Z) - H(Y) + H(Y) - H(X) && (23) \\ &= H(Z) - H(X) && (24) \\ &\approx 1.0143bit && (25) \end{aligned}$$

4.

$$\begin{aligned} I(Y;Z|X) &= H(Y) - H(X) && (26) \\ &\approx 0.6894bit && (27) \end{aligned}$$

5.

$$\begin{aligned} I(X;Z|Y) &= H(Z|Y) - H(Z|X,Y) && (28) \\ &= H(X) - H(X) && (29) \\ &= 0bit && (30) \end{aligned}$$

题目 2 这道题目似乎没有说明发送数字的先验概率分布，这里假设发送概率分布为均匀分布。设输入概率空间为 X，输出概率空间为 Y。

$$
\begin{aligned}
I(X;Y) &= \sum_x \sum_y P(x,y)\log_2 \frac{P(x|y)}{P(x)} && (31)\\
&= \sum_{x\in even(X)} \sum_y P(x,y)\log_2 \frac{P(x|y)}{P(x)} + \sum_{x\in odd(X)} \sum_{y\in odd(Y)} P(x,y)\log_2 \frac{P(x|y)}{P(x)} && (32)\\
&= \frac{1}{10} \sum_{x\in even(X)} \log_2 \frac{1}{P(x)} + \sum_{x\in odd(X)} \sum_{y\in odd(Y)} P(x,y)\log_2 \frac{P(x|y)}{P(x)} && (33)\\
&= \frac{1}{2}\log_2 10 + \sum_{x\in odd(X)} P(x,x)\log_2 \frac{P(x|x)}{P(x)} + \sum_{x\in odd(X)} \sum_{y\in odd(Y\setminus x)} P(x,y)\log_2 \frac{P(x|y)}{P(x)} && (34)\\
&= \frac{1}{2}\log_2 10 + \frac{5}{20}\log_2 5 + \sum_{x\in odd(X)} \sum_{y\in odd(Y\setminus x)} P(x,y)\log_2 \frac{P(x|y)}{P(x)} && (35)\\
&= \frac{1}{2}\log_2 10 + \frac{5}{20}\log_2 5 + \frac{20}{80}\log_2 \frac{10}{8} && (36)\\
&= \frac{1}{2}\log_2 10 + \frac{1}{4}\log_2 \frac{25}{4} && (37)\\
&\approx 2.3219bit && (38)
\end{aligned}
$$

那么问题来了，题目中要求计算“收到一个数字平均得到的信息量”，我不知道应该计算

$$I(X;Y), H(X,Y), H(X|Y), H(Y|X), H(X), H(Y)$$

中的哪一个。我这里的问题是不知道如何分析题目，来计算题目所要求的内容。以本题为例的话，就是题目中哪里透露了需要计算互信息而不是其他数据，请老师答疑解惑。

题目 3

1. 由下两小问的结果可得

$$
\begin{aligned}
H(Y,Z|X) &= H(Y|X) + H(Z|X,Y) && (39)\\
&= H(Y|X) + H(Z|X) - I(Y;Z|X) && (40)\\
&\le H(Y|X) + H(Z|X) && (41)
\end{aligned}
$$

   在 X 给定的条件下 Y, Z 相互独立时，等号成立。

2.

$$
\begin{aligned}
H(Y,Z|X) &= -\sum_x \sum_y \sum_z P(x,y,z)\log_2 P(y,z|x) && (42)\\
&= -\sum_x \sum_y \sum_z P(x,y,z)\log_2 (P(y|x)P(z|x,y)) && (43)\\
&= -\sum_x \sum_y \sum_z P(x,y,z)\log_2 P(y|x) - \sum_x \sum_y \sum_z P(x,y,z)\log_2 P(z|x,y) && (44)\\
&= -\sum_x \sum_y P(x,y)\log_2 P(y|x) - \sum_x \sum_y \sum_z P(x,y,z)\log_2 P(z|x,y) && (45)\\
&= H(Y|X) + H(Z|X,Y) && (46)
\end{aligned}
$$

3. 由互信息不等式可得

$$
\begin{aligned}
H(Z|X,Y) &= -\sum_x\sum_y\sum_z P(x,y,z)\log_2 P(z|x,y) && (47)\\
&= -\sum_x\sum_y\sum_z P(x,y,z)\log_2 (P(z|x,y)\frac{P(z|x)}{P(z|x)}) && (48)\\
&= -\sum_x\sum_y\sum_z P(x,y,z)\log_2 P(z|x) - \sum_x\sum_y\sum_z P(x,y,z)\frac{P(z|x,y)}{P(z|x)} && (49)\\
&= -\sum_x\sum_z P(x,z)\log_2 P(z|x) - \sum_x\sum_y\sum_z P(x,y,z)\frac{P(z|x,y)}{P(z|x)} && (50)\\
&= H(Z|X) - I(Y;Z|X) && (51)\\
&\le H(Z|X) && (52)
\end{aligned}
$$

题目 4　由凸函数的定义与对数和不等式可得

$$
\begin{aligned}
D(\lambda p_1+(1-\lambda)p_2||\lambda q_1+(1-\lambda)q_2) &= \sum_x(\lambda p_1(x)+(1-\lambda)p_2(x))\log_2\frac{\lambda p_1(x)+(1-\lambda)p_2(x)}{\lambda q_1(x)+(1-\lambda)q_2(x)} && (53)\\
&\le \sum_x(\lambda p_1(x)\log_2\frac{\lambda p_1(x)}{\lambda q_1(x)}+(1-\lambda)p_2(x)\log_2\frac{(1-\lambda)p_2(x)}{(1-\lambda)q_2(x)}) && (54)\\
&= \lambda D(p_1||q_1)+(1-\lambda)D(p_2||q_2) && (55)
\end{aligned}
$$