2022 秋季学期《应用信息论基础》作业 3

请于 2022.10.26 随堂提交，请写明姓名学号

1.设有二元离散无记忆信源，熵率为$H_\infty(U)$，试求，在序列长$N \to \infty$时

（1）不同典型序列的总数在全部可能序列中所占的比例

（2）信源序列中出现典型序列的概率

2. 设(X_i, Y_i)为 i.i.d$\sim p(x,y)$。假设X和Y独立与假设X和Y相关的对数似然比为

$$\frac{1}{n} log \frac{p(X^n)p(Y^n)}{p(X^n, Y^n)}$$

求$n \to \infty$时的极限。

3.设随机变量 X 取 4 个值，其概率分布为$\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{4}, \frac{1}{12}\right)$。

（1）请构造此随机变量的 Huffman 码

（2）证明存在两个不同的码字最优长度集，即证明码字长度分配(1,2,3,3)和(2,2,2,2)均是最优的

（3）由此可见，某些最优码的一些字符的相应码长有可能超过香农码的相应码长$\left\lceil log \frac{1}{p(x)} \right\rceil$。

4. 设有一个无记忆信源发出符号 A 和 B，已知$p(A) = \frac{1}{4}, p(B) = \frac{3}{4}$。

（1）针对该信源的二次扩展信源，给出香农-费诺编码方案并求编码后信息输出速率；

（2）针对该信源的三次扩展信源，给出霍夫曼编码方案并求编码后信息输出速率；

5. 设信源由一个离散随机变量 X表示，其取值范围是$A = \{\mathrm{a}_1, a_2, \dots, a_K\}$,熵为$H(X)$，若对该信源进行三元变长编码，其前缀码的平均码长为$\bar{l} = \frac{H(X)}{log3}$，试证明：

（1）对于每一个$a_k \in A, p(X = a_k) = 3^{-l_k}, k = 1,2, \dots, K$

（2）证明K为奇数

6. 对于等概信源

（1）证明：对于一个具有 n 个等概可能输出的信源，任意一个最优的前缀码的码字长度最多相差 1

（2）可变长度码的冗余被定义为$L-H$。一个具有 n 个等可能输出的随机变量，其中$2^m \le n \le 2^{m+1}$，对于最优二元可变长码，求出使得冗余$L-log_2 n$最大的 n 值。当$n\to\infty$时，在这种最坏的情形下，冗余的极限值（提示：0.0861）

7.信源编码定理表明，随机变量 X 的最优码的期望长度小于$H(X)+1$。请列举出一个随机变量，要求其最优码的期望长度近似等于$H(X)+1$，即对任意$\varepsilon>0$，试构造一个分布，使得其最优码的期望长度满足$L>H(X)+1-\varepsilon$