上周作业解答

Exercise 1.

Shannon 码需要已知概率质量函数 $p_i, 0 \le p_i \le 1, \sum_{i=1}^{M} p_i = 1$，其码长 $\ell_i = \lceil \log \frac{1}{p_i} \rceil$。若用了一个不匹配的概率质量函数 $q_i, 0 \le q_i \le 1$, $\sum_{i=1}^{M} q_i = 1$，则其码长是 $\tilde{\ell}_i = \lceil \log \frac{1}{q_i} \rceil$。

(1)求其平均码长，并计算不匹配概率对应的平均码长与匹配概率对应的平均码长之差。

(2)请设计一个程序例子，用 Matlab 程序仿真，体现上述计算的合理性。

上周作业解答

解答：

(1) 若使用了匹配的概率质量函数 p_i，可计算平均码长为

$$\sum_{i=1}^{M} p_i \lceil \log \frac{1}{p_i} \rceil$$

若使用了不匹配的概率质量函数 q_i，可计算平均码长为

$$\sum_{i=1}^{M} p_i \lceil \log \frac{1}{q_i} \rceil$$

它们之间的差 Δ 为

$$\sum_{i=1}^{M} p_i (\lceil \log \frac{1}{q_i} \rceil - \lceil \log \frac{1}{p_i} \rceil)$$

上周作业解答

(2) 仿真思路参考：给定一个确定的信源符号分布 $p_i, 0 \le p_i \le 1$, $\sum_{i=1}^{M} p_i = 1$，计算其平均码长。同时生成一个不匹配的分布 $q_i, 0 \le q_i \le 1, \sum_{i=1}^{M} q_i = 1$。依据Shannon码的构造规则，给出两种分布下每个符号的码长，计算其平均码长，并求差。

仿真可以通过由 M 种符号生成的文本序列进行，得到序列时，依据两种分布下进行编码，可以统计对序列编码的平均码长随着序列的长度的变化趋势。

上周作业解答

仿真取 $M = 10$，生成的 p 分布为:
[0.1215144 0.0072258 0.15319424 0.12132697 0.11715903 0.0920663 0.0570368 0.17259454 0.08351556 0.07436637]
计算相应平均码长为：3.7601 比特

仿真生成的 q 分布为:
[0.12317661 0.10493379 0.02668845 0.10184731 0.03657614 0.15573926 0.16935123 0.09801176 0.16788101 0.01579444]
计算相应平均码长为：4.3396 比特

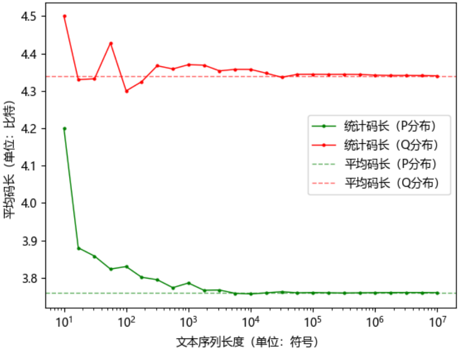

Figure: 平均码长与统计码长

上周作业解答

同时，观察统计码长之差的变化趋势

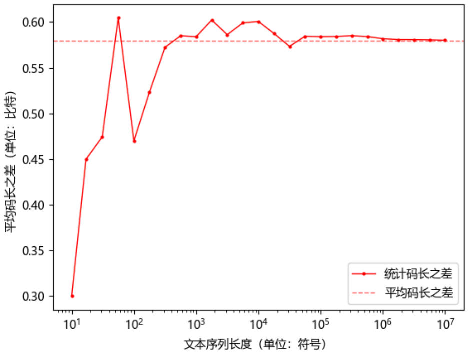

Figure: p 分布与 q 分布码长之差

不难发现，随着序列长度的增加，统计的平均码长趋于它们的期望值，因此它们的差也随着序列长度增加而趋向稳定。

Exercise 2.

记 $\mathcal{X}=\{0,1\}$ 为二元集，$\mathcal{X}^n$ 为所有 n 重二元数组。$x^n\in\mathcal{X}^n$ 的汉明重量是 x^n 中 1 的个数。

(1) $\mathcal{X}^n$ 中有多少个二元数组？

(2) 重量是 w 的 n 重二元数组有几个？

(3) 以 $n=4$ 为例，重为 2 的数组有 $(0,0,1,1)$，$(0,1,0,1)$，$(1,0,0,1)$，$(0,1,1,0)$，$(1,0,1,0)$，$(1,1,0,0)$ 6 个，因此我们称其可以承载 2 bit信息。即可以建立一个单射：$00\to(0,0,1,1)$，$01\to(0,1,0,1)$，$10\to(1,0,0,1)$，$11\to(0,1,1,0)$。记 $\mathcal{X}^n$中重量为 w 的全体二元数组为 T_w，求最大的 k，使得存在 $\mathcal{X}^k\to T_w$ 的单射。

(4) **思考题(不必完成)：** 任意给定 n,w，如何建立 $\mathcal{X}^k\to T_w$ 之间的单射？此单射可以称为等重编码，那么应该如何译码？

上周作业解答

解答：

(1) 一共有 2^n 个数组。

(2) 对于一个长为 n，重量为 w 的序列，可看作是在 n 个空位置中任意挑选 w 个，因此一共有 $\binom{n}{w}$ 种可能。

(3) 容易知道，$|T_w| = \binom{n}{w}$。只有当 $|\mathcal{X}^k| \leq |T_w|$ 时，存在 $\mathcal{X}^k \to T_w$ 的单射，此时

$$k \leq \log |T_w|$$

又因为 k 为整数，所以

$$k = \lfloor \log |T_w| \rfloor$$