QOJ12302 Unfair Card Deck

题目大意

有 m 种卡牌，总共 30 张，其中第 i 种卡牌有 a_i 张。保证 $1 \le a_i \le 2$。

有一个算法可以生成这 m 种卡牌的一个排列：它一个接一个地随机抽出一张卡牌，然后把它加入一个初始为空的序列的末尾，最终得到的序列就是生成的排列。

随机并不是均匀的。具体地，每种卡牌会有一个权值 X_i，记 $S=\sum c_i X_i$，其中 c_i 表示第 i 种卡牌剩余的数量，则每次会以 $\frac{c_i X_i}{S}$ 的概率抽中一张种类为 i 的卡牌。

给你该算法生成的 $n=100\,000$ 个序列，你要估计 X_i 的值。设你输出的解为 W_i，你只需要保证对于任意的 $X_i \le X_j$，有 $|\frac{X_i}{X_i+X_j}-\frac{W_i}{W_i+W_j}|<0.02$，且 $0<W_i \le 1$。

数据范围

- $n=100\,000$;
- $1 \le m \le 30$;
- $1 \le a_i \le 2$，$\sum a_i = 30$。

解题过程

考虑当 $a_i=1$ 时应该怎么做。

我们发现，卡牌 i 先于卡牌 j 出现的概率应当等于 $\frac{X_i}{X_i+X_j}$，这意味着，只有 **两张卡牌间的相对概率** 是重要的。我们可以用给出的 n 个序列，得到 i 先于 j 出现的频率，然后用频率估计概率。

如果我们知道了一个卡牌 i 的权值 X_i，那就可以从所有序列中删去 i，得到一个子问题，只需把子问题的解乘上权值 $1-X_i$。

我们发现卡牌 i 出现在第一个位置的概率应该等于 $\frac{X_i}{S}$。因此，可以直接使用 $\frac{o_i}{n}$（其中 o_i 是 i 出现在第一个位置的次数）估计卡牌 i 的概率。这可以保证一定的绝对误差，但无法保证相对误差。为了修复这一点，可以每次选出最大的 o_i 来估计 X_i，因为它具有足够多的样本。接下来只需删去 i 然后解决子问题即可。

上述为 $a_i=1$ 的解法。对于 $a_i=2$ 的那些卡牌，考虑在每个序列中只保留它们第一次出现的位置。转化后问题得到的解中 i 的权值乘了 2，所以最后需要把 W_i 除以 2。

参考资料

比赛题解