庄逸的数学与技术屋

最近播放列表的随机播放问题

Vortexer99

目录

1 问题情境

有这么一种音乐播放器，它有一个最近播放列表，其中总共能保存 $n(n \in \mathbb{Z}, n >= 2)$ 首歌。当选择某一首歌曲播放时，播放器将删除最近播放列表中第 n 首歌，并将第 1 至第 $n-1$ 首歌向后移一个排位，最后将选择播放的歌曲放入第一个排位，得到新的最近播放列表。注意，同一首歌曲能在最近播放列表中出现多次。

这个音乐播放器还有一种随机播放功能，即对一个有 n 首歌曲的播放列表，它能以每首歌 $1/n$ 的概率抽取其中一首播放。重复的歌曲是单独计算的，即如果一首歌出现了 m 次，总共就有 m/n 的概率播放这首歌。

我们的问题是，假设最初最近播放列表中有 n 首不同的歌曲，现在对最近播放列表进行随机播放操作。那么经过多长时间（即多少次播放）之后，最近播放列表中会仅剩下一首歌？每一首歌被剩下的概率又是多少？如果初始情况不是每首歌只出现一次，又会怎样呢？

2 一个可以解决具体数目问题的算法

2.1 想法

不妨考虑 $n=3$ 的情况，设最初列表中按顺序排着 ABC 三首歌，将 A 标记为 1，B 标记为 2，C 标记为 3，这个状态就记为 123. 再用 $N(123)$ 表示 123 状态达到平衡所需的步数期望，可以得到方程

$$N(123)=\frac{1}{3}(N(112)+1)+\frac{1}{3}(N(212)+1)+\frac{1}{3}(N(312)+1) \tag{1}$$

由等价性可知 $N(212)=N(121), N(312)=N(123)$, 那么

$$N(123)=\frac{1}{3}(N(112)+1)+\frac{1}{3}(N(121)+1)+\frac{1}{3}(N(123)+1) \tag{2}$$

对其中的各变元进行分析，同理可得

$$N(112)=\frac{2}{3}(N(111)+1)+\frac{1}{3}(N(122)+1) \tag{3}$$

$$N(121)=\frac{2}{3}(N(112)+1)+\frac{1}{3}(N(121)+1) \tag{4}$$

$$N(122)=\frac{1}{3}(N(112)+1)+\frac{2}{3}(N(121)+1) \tag{5}$$

注意到 $N(111)=0$, 整理得

$$\begin{aligned} N(123) &= \frac{1}{3}N(112)+\frac{1}{3}N(121)+\frac{1}{3}N(123)+1 && (6)\\ N(112) &= \frac{1}{3}N(122)+1 && (7)\\ N(121) &= \frac{2}{3}N(112)+\frac{1}{3}N(121)+1 && (8)\\ N(122) &= \frac{1}{3}N(112)+\frac{2}{3}N(121)+1 && (9) \end{aligned}$$

解得

$$N(122) = \frac{9}{2} \tag{10}$$

$$N(112) = \frac{5}{2} \tag{11}$$

$$N(121) = 4 \tag{12}$$

$$N(123) = \frac{19}{4} \tag{13}$$

2.2 编程实现

这个想法可以程序化实现，用 Mathematica 编程：

```
1  total = 5;
2  set = Tuples[Range[total], total];
3  eq = {};
4  var = {};
5  Print["initialize finished\n"];
6  Do[
7    a = set[[i]];
8    If[Length[Union[a]] == 1, AppendTo[eq, n[a] == 0],
9     AppendTo[eq,
10     n[a] == Total[
11       Table[1/total n[Prepend[Drop[a, −1], a[[k]]]] + 1/total, {k, 1,
12         total}]]]];
13   AppendTo[var, n[a]], {i, 1, Length@set}];
14 Print["equations generated\n"];
15 result = Solve[eq, var];
16 Print["equations solved, result is:"];
17 n[Range[total]] /. result[[1]]
18 (*其中前两行得到n个数所有可能的排列，也就是所有可能出现的播放列表状态。*)
19 (*6−13行计算出每种状态的方程，首先判断四个数是否一样，如果一样就加入形如n[{1,1,1,1}]==0的方
     程*)
20 (*如果不一样，就按照上面的想法计算出对应的方程并放入列表eq中*)
21 (*最后对方程组进行求解*)
```

最后可以得到准确值如下表

$amount$	2	3	4	5
N	2	$\frac{19}{4}$	$\frac{1179}{140}$	$\frac{12226997}{934830}$

2.3 关于概率

这个想法同样可以计算某首歌曲最后留下的概率，设 P 为最初标记的第一首歌曲留下的概率，只需将状态方程写为如

$$P(123) = \frac{1}{3}P(112) + \frac{1}{3}P(212) + \frac{1}{3}P(312) \tag{14}$$

的形式，也可解出方程组。同时，也得到其他歌曲留下的概率，如歌曲 2 留下的概率 $P_2(123) = P(213)$

2.4 小结

按照这个想法，理论上可以计算出特定 n 的所有情况的信息。但是事实上当 $n=6$ 时算期望就因过于复杂崩溃了。不过，从期望的准确值来看求其公式已经没有什么意义，我们可以只关心它的数量级。

不过，有趣的是经过计算，在列表 123 中歌曲 1 留下的概率为 $\frac{3}{6}$, 歌曲 2 留下的概率为 $\frac{2}{6}$，歌曲 3 留下的概率为 $\frac{1}{6}$

在列表 1234 中，同样有类似的规律，歌曲留下的概率依次为 $\frac{4}{10}\frac{3}{10}\frac{2}{10}\frac{1}{10}$. 可以推断，$n$ 首歌曲的列表中第 k 首歌曲留下的概率为

$$\frac{n+1-k}{n^2+n}$$

2.5 下一步工作

两个问题需要解决：一是达到稳定状态播放列表的播放次数期望的量级大小，二是为什么某一首歌曲留下的概率如上公式。

博客信息　此文章的博客来源：https://vortexer99.github.io/

自豪地采用 LaTeX！