

1 Problem

草原上有 2000 只黑羊和 2000 只白羊，它们不会死亡和新生。每天都有一只羊被随机选中，该随机选择与以前的选择无关。这只被选中的羊会找到一只与之异色的羊（如果还有的话）并对其鸣叫。一只黑羊对准一只白羊鸣叫会使白羊以概率 p 发生变色（变为黑羊）；一只白羊对准一只黑羊鸣叫会使黑羊以概率 q 发生变色（变为白羊）。管理者每天可以在神奇的羊变色现象发生前进场移走任意数量的白羊。管理者的目标是使得最后所剩黑羊数量的期望值尽可能大。譬如，头天就移走所有白羊，可以保证最后保住 2000 只黑羊。试针对一些不同的 p, q 取值来给管理者建议管理策略以及对最好可能的黑羊遗留数量做出估计。

2 Intuitive thoughts

假设某次操作完后，有 n 只黑羊和 m 只白羊时，那么此时黑羊变为白羊的概率是 $\frac{np}{n+m}$ ，白羊变为黑羊的概率是 $\frac{mq}{n+m}$ ，其余的 $\frac{n(1-p)+m(1-q)}{n+m}$ 什么都不会发生。这三个事件是互斥的，因此在 $\frac{np}{n+m} > \frac{mq}{n+m}$ 时，期望黑羊是变多的。因此，我以为每次将白羊移至这样的不等式成立即在期望下使黑羊变多。

3 Further thoughts

但是上述讨论并不完备，因为羊的数量是有限的。在无限的情况下，期望能够趋近取得，但是实际中并不是这样的——也许是风险非常大的：上述过程并没有保证黑羊的最小值。举例来说，我们抛一枚均匀的硬币，正面获得一块钱，反面失去一块钱，而玩家钱数有限，这样等价于 1d random walk，在无限步是有 1 的概率走到过 0 点的；而只要走到过 0，那么玩家就一无所有了。从这个角度来看，上述策略并不是最优策略——至少需要一些风险规避的手段。

注意到 $2000 + 2000$ 是个相对小的值，我们可以通过动态规划来解出最优值。设 $f_{i,j}$ 表示某天在移动羊之前有 i 只黑羊， j 只白羊，期望下能够在最后保留的黑羊的数量； $g_{i,j}$ 表示某天在移动羊之后有 i 只黑羊， j 只白羊，期望下能够在最后保留的黑羊的数量。那么有如下两个式子：

$$f_{i,j} = \max\{g_{i,k} | k \leq j\} \quad (1)$$

$$g_{i,j} = \left(\frac{i}{i+j} (pf_{i-1,j+1} + (1-p)f_{i,j}) + \frac{j}{i+j} (qf_{i+1,j-1} + (1-q)f_{i,j}) \right) \quad (2)$$

初始条件为

$$f_{i,0} = g_{i,0} = i \quad (3)$$

$$f_{0,i} = g_{0,i} = 0 \quad (4)$$

很不幸的是，上述动态规划转移是有环的，因此递推求解是困难的。但是我们可以使用迭代最短路的方式来进行迭代求解。注意到这是一个按 $i+j$ 分层的分层图，第 n 层只有 n 个需要迭代的变量，因此直观感受下在 $O(n^2)$ 次更新内应该能迭代出解，总更新次数应该在 $O(n^3)$ 左右，在现代计算机上大约需要几分钟的时间。我不确定这个复杂度对不对（毕竟迭代复杂度是玄学），但是可以试一试。

代码实现稍微有点复杂，写完后如果复杂度能对跑出结果的，我再向您报告结果好了。