Rat Exploration

Siyu Wang

Contents

1 Optimal Strategy

1.1 Discrete reward (two feeders always have different rewards)

Suppose all possible rewards are $r_1 < r_2 < \cdots < r_n$. Let r_a be the exploit reward $R_{exploit} = r_a$. For the explore reward, r_i has probability p_i to occur for $i \neq a$ and r_a has probability $p_a = 0$ to occur.

When the horizon is H, if to exploit, the expected reward is

$$EV(exploit) = r_a \cdot H,$$

if to explore, the expected value is

$$EV(explore) = \sum_{i=a+1}^{n} p_i \cdot r_i \cdot H + \sum_{i=1}^{a-1} p_i \cdot (r_i + (H-1) \cdot r_a) \tag{1}$$

$$\Delta EV = EV(explore) - EV(exploit) \tag{2}$$

$$= \sum_{i=a+1}^{n} p_i \cdot (r_i - r_a) \cdot H + \sum_{i=1}^{a-1} p_i \cdot (r_i - r_a) \tag{3}$$

ΔEV decreases as a function of a, see (3). Numerically, we can solve for a fixed list of r_i and p_i the optimal stopping threshold θ_H.

2 Notes

1. 80% of 5 drops vs 80% of 2 drops
2. always have X and X+ or - 2 drops, instead a random draw from 0,1,2,3,5
3. add penalty if the rat runs to a no-light feeder
4.