Unit 13: Mabinogion Sheep Problem

Instructor: Quan Zhou

13.1 Problem formulation

Mabinogion sheep problem. Consider a magical flock of sheep; some are black, and the others are white. At each time $n \in \mathbb{N} = \{1, 2, \dots\}$, a sheep is drawn randomly (with equal probability) from the whole flock and bleats. If the bleating sheep is white, one black sheep becomes white instantly; if the bleating sheep is black, a white sheep becomes black. Of course when all sheep are black or all are white, this magical process stops. Now suppose that we are allowed to do the following: at each time $n \in \mathbb{N}$, we can remove any number of white sheep from the flock. The Mabinogion sheep problem asks how to maximize the expected final number of black sheep.

Constructing the stochastic process of interest. Let's first develop a mathematical formulation of the problem. We use π_n to denote the number of white sheep we remove *before* the bleating at time n, and use W_n^π, B_n^π to denote the numbers of white and black sheep *after* the bleating at time n. The bivariate stochastic process $(W_n^\pi, B_n^\pi)_{n \geq 0}$ can be constructed as follows.

Let $(\Omega, \mathcal{F}, \mathsf{P})$ be a probability space on which we have independent random variables W_0, B_0 and $(U_n)_{n \geq 1}$ such that $U_1, U_2, \dots$ are i.i.d. with uniform distribution on $[0, 1]$. Let $W_0^\pi = W_0$, $B_0^\pi = B_0$. For each $n \geq 1$, we define $(W_n^\pi, B_n^\pi) = f(W_{n-1}^\pi, B_{n-1}^\pi, \pi_n, U_n)$ where

$$f(w, b, \pi, u) = \begin{cases} (w - \pi - 1, b + 1), & \text{if } (w - \pi) \wedge b > 0,\ u \leq \frac{b}{w - \pi + b}, \\ (w - \pi + 1, b - 1), & \text{if } (w - \pi) \wedge b > 0,\ u > \frac{b}{w - \pi + b}, \\ (w - \pi, b), & \text{if } (w - \pi) \wedge b \leq 0. \end{cases}$$

This implies that if $W_{n-1}^\pi - \pi_n > 0, B_{n-1}^\pi > 0$, then

$$\mathsf{P}(W_n^\pi = W_{n-1}^\pi - \pi_n + 1, B_n^\pi = B_{n-1}^\pi - 1 \mid W_{n-1}^\pi, B_{n-1}^\pi, \pi_n) = \frac{W_{n-1}^\pi - \pi_n}{W_{n-1}^\pi - \pi_n + B_{n-1}^\pi},$$
$$\mathsf{P}(W_n^\pi = W_{n-1}^\pi - \pi_n - 1, B_n^\pi = B_{n-1}^\pi + 1 \mid W_{n-1}^\pi, B_{n-1}^\pi, \pi_n) = \frac{B_{n-1}^\pi}{W_{n-1}^\pi - \pi_n + B_{n-1}^\pi},$$

which is the dynamics described in the Mabinogion sheep problem.

Admissible policies. We will also call $\pi = (\pi_n)_{n\geq 1}$ a policy, and we say a policy π is admissible if for each $n \geq 1$, (i) $\pi_n \in \{0, 1, 2, \dots, W^\pi_{n-1}\}$ a.s., and (ii) $\pi_n = \pi_n((W^\pi_0, B^\pi_0), \dots, (W^\pi_{n-1}, B^\pi_{n-1}))$; that is, π_n is measurable w.r.t. $\mathcal{F}^\pi_{n-1} = \sigma((W^\pi_k, B^\pi_k)_{0\leq k\leq n-1})$. Let Π denote the set of all admissible policies.

Remark 13.1. In property (ii) above, we can also replace $\mathcal{F}^\pi_{n-1}$ with a larger σ-algebra $\tilde{\mathcal{F}}_{n-1} = \sigma(W_0, B_0, U_1, U_2, \dots, U_{n-1})$. The latter is larger, since for any admissible π and $n \geq 1$, (W^π_n, B^π_n) is a measurable function of $W_0, B_0, U_1, U_2, \dots, U_n$. But note that given $(W^\pi_n, B^\pi_n)_{n\geq 0}$, we cannot exactly recover the sequence $(U_n)_{n\geq 1}$.

Objective. Our goal is to find $\max_{\pi\in\Pi} \mathsf{E}[B^\pi_\infty]$, where $B^\pi_\infty = \lim_{n\to\infty} B^\pi_n$ denotes the final number of black sheep.

Lemma 13.1. *Suppose $W_0 = w_0, B_0 = b_0$ a.s. for some $w_0, b_0 \geq 0$. For any $\pi \in \Pi$, let $T^\pi = \inf\{n \geq 0\colon \min\{W^\pi_n, B^\pi_n\} = 0\}$. Then, $\mathsf{E}[T^\pi] < \infty$, and thus $B^\pi_\infty = \lim_{n\to\infty} B^\pi_n$ exists.*

Proof. Let $N = W_0 + B_0$ and fix some $\pi \in \Pi$. As long as $B^\pi_{n-1} > 0$, we have

$$\frac{B^\pi_{n-1}}{W^\pi_{n-1} - \pi_n + B^\pi_{n-1}} \geq \frac{1}{N},$$

since $\pi_n \geq 0$ and the total number of sheep cannot increase. This implies that $\mathsf{P}(T^\pi \leq n+N \mid \mathcal{F}^\pi_n) \geq N^{-N}$. Hence, $\mathsf{E}[T^\pi] < \infty$, which implies $T^\pi < \infty$, a.s. The asserted result follows since B^π_∞ exists on the event $\{T^\pi < \infty\}$. $\square$

Value function. It will be convenient to solve this problem for all possible initial states simultaneously. Let $\mathbb{N}_0 = \{0, 1, 2, \dots\}$. For any $(w, b) \in \mathbb{N}_0^2$, we can define a new probability measure on $(\Omega, \mathcal{F})$ (assuming that it is sufficiently rich) such that $W^\pi_0 = w, B^\pi_0 = b$, a.s. and the distribution of $(U_n)_{n\geq 1}$ remains the same. Denote this probability measure by $\mathsf{P}_{w,b}$ and the corresponding expectation by $\mathsf{E}_{w,b}$. Define

$$V(w, b) = \max_{\pi\in\Pi} \mathsf{E}_{w,b}[B^\pi_\infty]. \tag{1}$$

We want to find the expression for $V(w, b)$.

13.2 Optimal policy

Solution. We claim that the optimal policy is given by $\pi^*_n = g(W^*_{n-1}, B^*_{n-1})$, where $(W^*_n, B^*_n) = (W^{\pi^*}_n, B^{\pi^*}_n)$ denotes the optimally controlled process and the measurable function g is defined by

$$g(w,b) = \begin{cases} \max\{0, w-b+1\}, & \text{if } w>0, b>0, \\ 0, & \text{otherwise.} \end{cases}$$

Under the policy π^*, once there is no black or white sheep, the whole process stops (i.e., the values of W^*_n, B^*_n no longer change). And if there are more white sheep than black sheep (assuming there is at least one black sheep), we reduce the number of white sheep to that of black sheep minus one. The probability that a black sheep bleats at time n is given by $p(W^*_{n-1}, B^*_{n-1})$ where the measurable function p is defined by

$$p(w,b) = \frac{b}{w - g(w,b) + b}.$$

Now define

$$v(w,b) = \mathsf{E}_{w,b}[B^*_\infty]. \tag{2}$$

To prove the optimality of π^* is equivalent to showing that $v = V$, where V is defined in (1). The next lemma describes how to find v.

Lemma 13.2. *The function $v\colon \mathbb{N}_0^2 \to [0,\infty)$ defined in (2) is the unique solution to the following system of equations:*

(i) $v(0,b) = b$ *for any* $b \in \mathbb{N}$.

(ii) $v(w,b) = v(w-1,b)$ *for any* (w,b) *such that* $w \geq b$ *and* $w > 0$.

(iii) $v(w,b) = \frac{w}{w+b} v(w+1, b-1) + \frac{b}{w+b} v(w-1, b+1)$ *for any* (w,b) *such that* $w < b$, $b > 0$ *and* $w > 0$.

Proof. Condition (i) is obvious, since the whole process stops at time 0 if $W^*_0 = 0$. To prove conditions (ii) and (ii), we will use a Markov chain argument. Write $\mathcal{F}^*_n = \mathcal{F}^{\pi^*}_n$. Observe that under the policy π^*, $(W^*_n, B^*_n)_{n\geq 0}$ is a bivariate Markov chain; that is, the conditional distribution of (W^*_n, B^*_n) given $\mathcal{F}^*_{n-1}$ is the same as the conditonal distribution of (W^*_n, B^*_n) given (W^*_{n-1}, B^*_{n-1}). More explicitly, given $W^*_{n-1} = w, B^*_{n-1} = b$,

- if $w = 0$ or $b \leq 1$, then $W_n^* = w - g(w,b)$ and $B_n^* = b$;
- if $w \geq 1$ and $b \geq 2$, then with probability $1 - p(w,b)$, we have $W_n^* = w - g(w,b) + 1, B_n^* = b - 1$; with probability $p(w,b)$, we have $W_n^* = w - g(w,b) - 1, B_n^* = b + 1$.

In the first case where $w = 0$ or $b \leq 1$, the process stops at either time 0 (if $w = 0$ or $b = 0$) or time 1 (if $w \geq 1, b = 1$), and thus $v(w,b) = b$. To characterize v in the second case, we use standard results from Markov chain theory (which we do not prove here) to get

$$\mathsf{E}\left[\lim_{m \geq n} B_m^* \,\middle|\, W_n^* = w, B_n^* = b\right] = \mathsf{E}_{w,b}\left[\lim_{m \geq 0} B_m^*\right] = v(w,b)$$

whenever $\mathsf{P}(W_n^* = w, B_n^* = b) > 0$. Hence, by applying the law of total expectation and conditioning on W_1^*, B_1^*, we find that if $w \geq 1$ and $b \geq 2$,

$$\begin{aligned} v(w,b) = (1 - p(w,b))\, v(w - g(w,b) + 1, b - 1) + \\ p(w,b)\, v(w - g(w,b) - 1, b + 1). \end{aligned} \tag{3}$$

Now consider condition (ii). It clearly holds if $b = 0$ and $w > 0$ ($v = 0$ in this case). If $w \geq b > 0$, we have $w - g(w,b) = b - 1$. By (3),

$$v(w,b) = \frac{b-1}{2b-1} v(b, b-1) + \frac{b}{2b-1} v(b-2, b+1) = v(b-1, b).$$

This proves condition (ii). For condition (iii), observe that in this case we have $g(w,b) = 0$, and then condition (iii) follows from (3).

We leave it as an exercise to show that conditions (i), (ii), (iii) uniquely determine v. □

To elucidate the line of reasoning behind our proof of the optimality of π^*, we will first prove a theorem by assuming that the function v has a certain property. Later we will show that this assumption always holds.

Theorem 13.1. *Suppose that $X_n^\pi = v(W_n^\pi, B_n^\pi)$ is a supermartingale w.r.t. $(\mathcal{F}_n^\pi)_{n \geq 0}$ for any $\pi \in \Pi$, where v is as given in Lemma 13.2. Then, $v = V$, where V is defined in* (1).

Proof. Fix arbitrary w, b and let $W_0 = w, B_0 = b$. By Lemma 13.2, B^π_∞ exists, and an analogous argument shows that W^π_∞ also exists. Since W^π_n, B^π_n take values in $\mathbb{N}$, we have

$$X^\pi_\infty = \lim_{n\to\infty} X^\pi_n = v(W^\pi_\infty, B^\pi_\infty) = B^\pi_\infty,$$

where in the last equality we have used $\min\{W^\pi_\infty, B^\pi_\infty\} = 0$, which holds by Lemma 13.2. By (2), v should be non-negative. Hence, Corollary 5.1 implies that $\mathsf{E}[B^\pi_\infty] = \mathsf{E}[X^\pi_\infty] \le X^\pi_0 = v(w,b)$. Since $v(w,b) = \mathsf{E}_{w,b}[B^*_\pi]$, this shows that v coincides with V. $\square$

Lemma 13.3. *The assumption in Theorem 13.1 holds; i.e., $X^\pi_n = v(W^\pi_n, B^\pi_n)$ is a supermartingale w.r.t. $(\mathcal{F}^\pi_n)_{n\ge 0}$ for any $\pi \in \Pi$.*

Proof. We prove a slightly stronger result: $(X^\pi_n)_{n\ge 0}$ is a supermartingale w.r.t. $(\tilde{\mathcal{F}}_n)_{n\ge 0}$, where $\tilde{\mathcal{F}}_n = \sigma(W_0, B_0, U_1, \dots, U_n)$ (recall Remark 13.1). We claim that it suffices to show that

(a) $v(w,b) \ge v(w-1,b)$ if $w > 0$;

(b) $v(w,b) \ge \frac{w}{w+b} v(w+1, b-1) + \frac{b}{w+b} v(w-1, b+1)$ if $w > 0, b > 0$.

To see this, recall that we can write $(W^\pi_n, B^\pi_n) = f(W^\pi_{n-1}, B^\pi_{n-1}, \pi_n, U_n)$ for some measurable function f. Hence, $X^\pi_n = (v \circ f)(W^\pi_{n-1}, B^\pi_{n-1}, \pi_n, U_n)$. It follows that

$$\mathsf{E}[X^\pi_n \,|\, \tilde{\mathcal{F}}_{n-1}] = \mathsf{E}\left[X^\pi_n \,|\, W^\pi_{n-1}, B^\pi_{n-1}, \pi_n\right],$$

since U_n is independent of $\tilde{\mathcal{F}}_{n-1}$. If $W^\pi_{n-1} - \pi_n = 0$ or $B^\pi_{n-1} = 0$, the process stops at time n before the bleating, which yields

$$\begin{aligned}\mathsf{E}\left[X^\pi_n \,|\, W^\pi_{n-1}, B^\pi_{n-1}, \pi_n\right] &= B^\pi_{n-1} \\ = v(W^\pi_{n-1} - \pi_n, B^\pi_{n-1}) &\le v(W^\pi_{n-1}, B^\pi_{n-1}) = X^\pi_{n-1}\end{aligned}$$

where the inequality follows from condition (a). If $W^\pi_{n-1} - \pi_n > 0$ and $B^\pi_{n-1} > 0$, we use conditions (a) and (b) to get

$$\begin{aligned}&\mathsf{E}\left[X^\pi_n \,|\, W^\pi_{n-1} = w, B^\pi_{n-1} = b, \pi_n = \pi\right] \\ &= \frac{w-\pi}{w-\pi+b} v(w-\pi+1, b-1) + \frac{b}{w-\pi+b} v(w-\pi-1, b+1) \\ &\le v(w-\pi, b) \\ &\le v(w,b) = X^\pi_{n-1}.\end{aligned}$$

Proof of conditions (a) and (b) is left as an exercise. $\square$

Exercise 13.1. Find $v(10, 10)$ for v given in Lemma 13.2. (Of course, you can write computer code to calculate this.)

Exercise 13.2. Complete the proof of Lemma 13.3; that is, prove conditions (a) and (b).

Exercise 13.3. Let $Z_1, Z_2, \dots$ be i.i.d. with $\mathsf{P}(Z_1 = 1) = p$ and $\mathsf{P}(Z_1 = -1) = 1 - p =: q$, and let $\mathcal{F}_n = \sigma(Z_1, \dots, Z_n)$. We interpret Z_n as your winnings per unit stake in the n-th game. Let $S_0 > 0$ represent your initial balance, and π_n be your stake in the n-th game. We say $\pi = (\pi_n)_{n \geq 1}$ is admissible if (i) π is previsible w.r.t. $(\mathcal{F}_n)_{n \geq 0}$, and (ii) $0 \leq \pi_n \leq S^\pi_{n-1}$ for each n, where S^π_n denotes your balance after the n-th game. That is, the dynamics of $(S_n)_{n \geq 0}$ is given by

$$S^\pi_n = S^\pi_{n-1} + \pi_n Z_n.$$

Given some fixed $N \in \{1, 2, \dots\}$, we want to find the optimal strategy π that maximizes $\mathsf{E}[\log(S^\pi_N / S_0)]$. Solve the following questions.

(i) For any $s > 0$, find the maximum of $\mathsf{E}[\log S^\pi_n \mid S^\pi_{n-1} = s]$ over all admissible choices of π_n.

(ii) Show that for any admissible π, $Y^\pi_n = \log S^\pi_n - n\alpha$ is a supermartingale, where $\alpha = p \log p + q \log q + \log 2$.

(iii) Find the admissible π that maximizes $\mathsf{E}[\log(S^\pi_N / S_0)]$ and prove your claim.

References

[1] David Williams. *Probability with martingales*. Cambridge university press, 1991.