Doob's Optional Stopping Theorem and the ABRACADABRA Problem

Yi Gu

June 24, 2020

This is a solution to Exercise 10.6 in the book, and the exercise is about a monkey typing random letters.

1 The problem

Exercise 1. *At each of times* $1, 2, 3, \ldots,$ *a monkey types a capital letter at random, the sequence of letters typed forming an IID sequence of random variables each chosen uniformly from amongst the* 26 *possible capital letters.*

Just before each time $n = 1, 2, \ldots,$ *a new gambler arrives, He bets* \$1 *that the* n*-th letter will be* A.

If he loses, he leaves. If he wins, he receives \$26 *all of which he bets on the event that the* $(n+1)$*-th letter will be* B.

If he loses, he leaves. If he wins, he bets his whole fortune of \$$26^2$ *that the* $(n+2)$*-th letter will be* R.

And so on through the ABRACADABRA sequence.

Let T *be the first time by which the monkey has produced the consecutive sequence ABRACADABRA. Show that*

$$E(T) = 26^1 1 + 26^4 + 26$$

To solve this, a naive thought is that ABRACADABRA is 11 lettes long, and the probability of a random 11-letter word being ABRACADABRA is exactly $(\frac{1}{26})^{11}$, and if type 11 letters is one trial, the expected number of trials is 26^{11}.

However, this idea broke up the random string into 11-letter substrings, while this word may start in the middle of a block. In other words, we only consider strings whose starting position is divisible by 11.

2 An intuitive idea

Suppose we open a casino, just next to the monkey and the typewriter.

Before each keystroke, a gambler comes to our casino, and gamble in the same way as described in the problem statement.

Suppose our casino is fair, i.e. the expected outcome is zero, which means if the gambler bets \$1, he will receive \$26 if he wins.

We keep our casino until the monkey types ABRACADABRA for the first time, and denote the number of keystrokes up to this time as T. Our revenue so far will be exactly T dollars.

How much shall we pay for the gamblers? There is one gambler coming in just before the last keystroke, who wins \$26. Also there is one gambler winning his 4 first bets ABRA, worthing $\$26^4$. Finally, the biggest winner wins $\$26^{11}$. No other gambler should be paid. Thus our casino will have to pay the gamblers $26^{11} + 26^4 + 26$ dollars.

Remember we are so kind that our casino is fair. Thus we have

$$E(T) = 26^{11} + 26^4 + 26.$$

3 A formal proof

Let $(U_i)_{i\in\mathbb{N}}$ denote random letters drawn independently and uniformly from the English alphabet, i.e. $(U_i)_{i\in\mathbb{N}}$ are i.i.d. random variables, uniformly distributes in the set $\{\text{A}, \dots, \text{Z}\}$.

Define T as $T := \min\{n \in \mathbb{N}, n \geq 11 : U_{[n-10,n]} = \text{ABRACADABRA}\}$.

Let $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be the natural filtration of $(U_i)_{i\in\mathbb{N}}$. We are going to prove the following results.

Proposition 2. *T is a stopping times with $E(T) < \infty$.*

Proposition 3. *There exists a martingale $M = (M_n)_{n\in\mathbb{N}_0}$ such that:*

1. *$M_0 = 0$, $M_T = 26^1 1 + 26^4 + 26 - T$;*
2. *$\exists C \in (0, \infty)$ such that $|M_{n+1} - M_n| \leq C$, $\forall n \in \mathbb{N}_0$.*

Combining these two proposition, one obtains immediately the result of the exercise by Doob's optional sampling theorem.

Proof of proposition 2. $\{T \leq n\} = \bigcup_{i=11}^{n} \{U_{[i-10,i]} = \text{ABRACADABRA}\} \in \mathcal{F}_n$. So T is a stopping time.

Let event A_n defined as

$$A_n := \{U_{[n+1,n+11]} = \text{ABRACADABRA}\}.$$

As U_i are independently and uniformly chosen letters, $P(A_n) = p^{11} > 0$ where $p = \dfrac{1}{26}$.

Since $A_n \subseteq \{T \leq n + 11\}$, we have $P(T \leq n + 11 \mid T > n) \geq P(A_n \mid T > n)$. Note that A_n and $\{T > n\}$ are independent, so $P(T \leq n + 11 \mid T > n) \geq P(A_n) = p^{11}$, which is sufficient to have $E(T) < \infty$. □

Proof of proposition 3. Denote by x_i the i-th letter in ABRACADABRA. The capital of the j-th gambler at time n is given by the process $(K_n^j)_{n\in\mathbb{N}_0}$, where

$$K_n^j = \begin{cases} 1 & n < j \\ K_{n-1} \cdot 26\mathbf{1}_{U_n = x_n} & j \leq n \leq j + 10 \\ K_{11} & n > j + 10 \end{cases}.$$

And we define

$$M_0 = 0, \quad M_n = \sum_{j=1}^{n} (K_n^j - K_0^j) = (\sum_{j=1}^{n} K_n^j) - n.$$

We have $M_T = 26^{11} + 26^4 + 26 - T$, as $K_T^{T-10} = 26^{11}$, $K_T^{T-3} = 26^4$ and $K_T^T = 26$.

For every fixed j, $(K_n^j)_{n \in \mathbb{N}_0}$ is a martingale as

$$E[K_n^j \mid \mathcal{F}_{n-1}] = K_{n-1}^j \cdot 26 P(U_n = x_{n-j+1}) = K_{n-1}^j,$$

because U_n is independent of $\mathcal{F}_{n-1}$.

As

$$E[M_n \mid \mathcal{F}_{n-1}] = \sum_{j=1}^{n} E[K_n^j \mid \mathcal{F}_{n-1}] - n = \sum_{n-1}^{j} -n = \sum j = 1^{n-1} K_{n-1}^j + 1 - n = M_{n-1},$$

M is a martingale.

Also not that $|K_n^j - K_{n-1}^j| \leq 25^{11}$, and $|M_n - M_{n-1}| \leq \sum_{j=n-10}^{n} |K_n^j - K_{n-1}^j| \leq 11 \cdot 25^{11}$, i.e. M has bounded increments. $\square$