COIN TOSSING EXERCISE

IORDAN GANEV

1. The Problem

Alice tosses a fair coin until the sequence HTH appears, then stops. Meanwhile, Bob tosses another fair coin until HTT appears, then stops. On average, how many tosses before each of them stops?

2. The Solution

We first introduce notation. Let X_i be the random variable recording the outcome of the i-th coin toss. Thus, X_i takes values in $\{H, T\}$, each with probability $1/2$. We define the following random variables:

$$A = \min\{i \ : \ X_{i-2}X_{i-1}X_i = HTH\}$$
$$B = \min\{i \ : \ X_{i-2}X_{i-1}X_i = HTT\}$$

The random variables A and B take values in the integers $\mathbb{Z}$, although $P(A = k) = P(B = k) = 0$ if $k < 3$. We seek $E(A)$ and $E(B)$.

The idea is to consider the outcome of the first few toss, and to express the expected value of A and B in terms of their expected values given certain starting subsequences. If the first toss is tails (i.e. $X_1 = T$), then the process for both counting procedures restarts in the sense that

$$P(A = k \mid X_1 = T) = P(A = k - 1) \qquad P(B = k \mid X_1 = T) = P(B = k - 1). \tag{2.1}$$

It follows (e.g. from the definition of expected value) that

$$E(A \mid X_1 = T) = E(A) + 1 \qquad E(B \mid X_1 = T) = E(B) + 1 \tag{2.2}$$

An informal way to think about these equations is as follows. We are looking for the sequences HTH and HTT. So, when an H appears, we are alert since it may be the start of HTH or HTT. On the other hand, a T at the first toss means that there's no hope of obtaining HTH in the first three tosses, so it is as if our experiment starts over, with the accrued cost of one toss.

Using elementary identities (see appendix), we compute:

$$\begin{aligned} E(A) &= E(A \mid X_1 = H)P(X_1 = H) + E(A \mid X_1 = T)P(X_1 = T) \\ &= \frac{1}{2}E(A \mid X_1 = H) + \frac{1}{2}(E(A) + 1) \end{aligned}$$

The same computation holds for B, and we obtain:

$$E(A) = E(A \mid X_1 = H) + 1 \qquad \text{and} \qquad E(B) = E(B \mid X_1 = H) + 1$$

We do a similar analysis for $E(A \mid X_1 = H)$ and for $E(B \mid X_1 = H)$:

$$\begin{aligned} E(A \mid X_1 = H) &= E(A \mid X_1X_2 = HH)P(X_2 = H) \\ &\quad + E(A \mid X_1X_2X_3 = HTH)P(X_2X_3 = TH) \\ &\quad + E(A \mid X_1X_2X_3 = HTT)P(X_2X_3 = TT) \\ &= \frac{1}{2}[E(A \mid X_1 = H) + 1] + \frac{1}{4}(3) + \frac{1}{4}[E(A) + 3] \end{aligned}$$

$$\begin{aligned} E(B \mid X_1 = H) &= E(B \mid X_1X_2 = HH)P(X_2 = H) \\ &\quad + E(B \mid X_1X_2X_3 = HTH)P(X_2X_3 = TH) \\ &\quad + E(B \mid X_1X_2X_3 = HTT)P(X_2X_3 = TT) \\ &= \frac{1}{2}[E(B \mid X_1 = H) + 1] + \frac{1}{4}[E(B \mid X_1 = H) + 2] + \frac{1}{4}(3) \end{aligned}$$

It follows that $2E(A \mid X_1 = H) = E(A) + 8$ and $E(B \mid X_1 = H) = 7$. Combining with Equation 2.2, we obtain

$$E(A) = 10 \qquad \text{and} \qquad E(B) = 8.$$

3. Appendix

Let Ω be a sample space. Suppose $\Omega = \coprod_i \Omega_i$ is a disjoint union of subsets Ω_i. Let $\Theta \subset \Omega$ be a subset. One can easily verify the following identities:

(1) $P(\Theta) = \sum_i P(\Theta \cap \Omega_i)$
(2) $P(\Theta) = \sum_i P(\Theta \mid \Omega_i)P(\Omega_i)$
(3) $\sum_i P(\Theta|\Omega_i) = 1$

Let X be a random variable defined on Ω that takes values in the integers $\mathbb{Z}$. Then, for any integer m,

(1) $E(X) = \sum_i E(X|\Omega_i)P(\Omega_i)$
(2) $\sum_k kP(X = (k - m)) = E(X) + m$