We begin with an illustrative problem.

> **Problem.** What is the expected number of hits taken for a monkey to type out the word ABRACADABRA on a typewriter containing only the 26 letters A-Z, given that a random letter is depressed on each hit? (Adapted from Williams' *Probability with Martingales.*)

Naively we might imagine the answer to be 11×26^{11}, because there are 11 letters in ABRACADABRA, and each letter has a 1-in-26 probability of getting chosen. If each 11-letter attempt of typing out ABRACADABRA were independent, then the number of attempts X required for the first success follows a geometric distribution with parameter $p = 1/26^{11}$, and $\langle X \rangle = p^{-1} = 26^{11}$. We then multiply this result by 11 because each attempt consists of 11 letters.

But this answer is incorrect—successive trials here are not independent. Say we have the failed attempt XXXXXABRACA. A success on the subsequent attempt requires only the 5-letter combination DABRA, which clearly increases the probability of success. Our naive answer of 11×26^{11} must therefore be greater than the actual answer.

The lack of independence makes this problem non-trivial, but an elegant solution arises from the study of martingales. A stochastic process $(M_n;\, n \geq 0)$ is a *martingale* if

$$\langle M_{n+1} \mid M_n, M_{n-1}, \ldots, M_0 \rangle = M_n$$

and $\langle M_n \rangle$ exists for all n. In the context of gambling, from which the term martingale is derived, a martingale describes the result of a fair game: let M_n be the amount of money a certain player has at turn n. His expected amount of money at turn $n+1$, given complete information about the past history of the game, is equal to M_n—on average, the player should neither gain nor lose money, and the game is fair.

The martingale solution then follows by constructing a martingale related to the problem and using the optional stopping theorem, but this development is too lengthy to pursue. Rather, recalling the connection between a martingale and a fair game, we will construct a fair game to solve the problem.

Let a gambler bet \$1 on each letter generated by the monkey. If that letter is A, he wins \$26 and must bet it all on the next letter. If the next letter in the sequence is B, he wins $\$26^2$ and must bet it all on the next letter, and so on and so forth until the word ABRACADABRA is spelled out. This game is fair: at each stage the person has a 1/26 chance of winning 26 times what he bets. Let T be the expected number of hits before the monkey spells out ABRACADABRA. In the context of the game, on average, when ABRACADABRA is first spelled out, the gambler will have won $\$26^{11} + \$26^4 + \$26$ (from the sequences ABRACADABRA, ABRA, and A respectively) and lost $\$T$. Because the game is fair, the expected profit is zero. Hence

$$0 = 26^{11} + 26^4 + 26 - T.$$

(Interestingly, we see that the expected number of hits goes up due to the repeating sequence ABRA, which appears both at the start and at the end of ABRACADABRA, and to the repeating letter A, which appears likewise. Why should this be the case? With regards to ABRA, let us imagine that ABRACADXXXX were the desired sequence. Then ABRACADABRA would be a failed sequence, but one in which the last four letters could potentially be the first four letters of ABRACADXXXX. There is no equivalent when the desired sequence is ABRACADABRA, so the expected number of hits should be greater for this latter sequence. The same is true with regards to the repeating letter A.)

We have shown how constructing fair games can help us solve problems commonly associated with martingales, the underlying concepts being the same in both cases. The fair-game construction is more intuitive to visualize and construct a fair game than the corresponding martingale, which is far from being immediately obvious. Nevertheless, martingale theory is well developed, and the identification of a suitable martingale and subsequent analysis will likely provide more complete information about a system than the construction of a fair game.

> **Exercise.** Consider a one-dimensional random walk on the integers starting at the origin, which moves one step right with probability p and one step left with probability $1 - p$ and stops when it reaches a or b, $a < 0 < b$. Find the probability that the random walk reaches b.