17 Three Applications

Parrondo’s Paradox

This famous paradox was constructed by Spanish physicist J. Parrondo. We will consider three games $A$, $B$ and $C$ with five parameters: probabilities $p$, $p_1$, $p_2$, and $\gamma$, and an integer period $M \geq 2$. These parameters are, for now, general so that description of the games is more transparent. We will choose particular values when we are done with the analysis.

We will call a game losing if after playing it for a long time a player’s capital becomes more and more negative, i.e., the player loses more and more money.

The game $A$ is very simple; in fact it is an asymmetric one-dimensional simple random walk. Win $1$, i.e. add $+1$ to your capital, with probability $p$ and lose a dollar, i.e. add $-1$ to your capital, with probability $1-p$. This is clearly a losing game if $p < \frac{1}{2}$.

In game $B$, the winning probabilities depend on whether your current capital is divisible by $M$. If it is, you add $+1$ with probability $p_1$, and $-1$ with probability $1-p_1$, and if it is not, you add $+1$ with probability $p_2$ and $-1$ with probability $1-p_2$. We will determine when this is a losing game below.

Now consider game $C$, in which you, at every step, play $A$ with probability $\gamma$ and $B$ with probability $1-\gamma$. Is it possible that $A$ and $B$ are losing games, while $C$ is winning?

The surprising answer is yes! However, this should not be so surprising as in game $B$ your winning probabilities depend on the capital you have and you can manipulate the proportion of time your capital spends at “unfavorable” amounts by playing the combination of the two games.

Now on to the analysis. As mentioned, game $A$ is easy. To analyze game $B$, take a simple random walk which makes a $+1$ step with probability $p_2$ and $-1$ step with probability $1-p_2$. Assume that you start this walk from some $x$, $0 < x < M$. Then, by the Gambler’s ruin computation (Example 11.6),

$$ P(\text{the walk hits } M \text{ before } 0) = \frac{1 - \left(\frac{1-p_2}{p_2}\right)^x}{1 - \left(\frac{1-p_2}{p_2}\right)^M}. $$

Starting from a multiple of $M$, the probability that you increase your capital by $M$ before either decreasing it by $M$ or returning to the starting point is

$$ p_1 \cdot \frac{1 - \left(\frac{1-p_2}{p_2}\right)^{M}}{1 - \left(\frac{1-p_2}{p_2}\right)^M}. $$

(You have to make a step to the right, and then use the formula (17.1) with $x = 1$.) Similarly, from a multiple of $M$, the probability that you decrease your capital by $M$ before either increasing
it by $M$ or returning to the starting point is

$$(17.3) \quad (1 - p_1) \cdot \frac{\left(\frac{1-p_2}{p_2}\right)^{M-1} - \left(\frac{1-p_2}{p_2}\right)^M}{1 - \left(\frac{1-p_2}{p_2}\right)^M}.$$  

(Now you have to move one step to the left and then use $1-(\text{probability in (17.1) with } x = M - 1)$.)

The main trick is to observe that game $B$ is losing if $(17.2) < (17.3)$. Why? Observe your capital at multiples of $M$: if the probability that you get to the lower multiple is higher than the probability that you get to the higher multiple, then the game is losing, and that is exactly when $(17.2) < (17.3)$. After some algebra, this condition reduces to

$$(17.4) \quad \frac{(1 - p_1)(1 - p_2)^{M-1}}{p_1 p_2^{M-1}} > 1.$$  

Now game $C$ is the same as game $B$, with $p_1$ and $p_2$ replaced by $q_1 = \gamma p + (1 - \gamma)p_1$ and $q_2 = \gamma p + (1 - \gamma)p_2$, obtaining a winning game if

$$(17.5) \quad \frac{(1 - q_1)(1 - q_2)^{M-1}}{q_1 q_2^{M-1}} < 1.$$  

This is easily achieved with large enough $M$ as soon as $p_2 < \frac{1}{2}$ and $q_2 > \frac{1}{2}$, but even for $M = 3$, one can choose $p = \frac{5}{11}$, $p_1 = \frac{1}{11}$, $p_2 = \frac{10}{11}$, $\gamma = \frac{1}{2}$, to get $\frac{6}{5}$ in (17.4) and $\frac{24}{300}$ in (17.5).

A Discrete Renewal Theorem

**Theorem 17.1.** Assume that $f_1, \ldots, f_N \geq 0$ are given numbers, with $\sum_{k=1}^N f_k = 1$. Let $\mu = \sum_{k=1}^N k f_k$. Define the sequence $u_n$ as follows: $u_n = 0$ if $n < 0$, $u_0 = 1$, $u_n = \sum_{k=1}^N f_k u_{n-k}$ if $n > 0$.

Assume that the greatest common divisor of the set $\{k : f_k > 0\}$ is 1. Then

$$\lim_{n \to \infty} u_n = \frac{1}{\mu}.$$  

**Example 17.1.** Roll a fair die forever and let $S_m$ be the sum of outcomes of first $m$ rolls. Let $p_n = P(S_m \text{ ever equals } n)$. Estimate $p_{10,000}$.  

One can write a linear recursion

\[ p_0 = 1, \]
\[ p_n = \frac{1}{6} (p_{n-1} + \cdots + p_{n-6}), \]

and then solve it, but this is a lot of work! (Note that one should either modify the recursion for \( n \leq 5 \) or, much easier, define \( p_n = 0 \) for \( n < 0 \).) By the above theorem, however, we can immediately conclude that \( p_n \) converges to \( \frac{2}{7} \).

**Example 17.2.** Assume a random walk starts from 0 and jumps from \( x \) either to \( x + 1 \) or to \( x + 2 \), with probability \( p \) and \( 1 - p \), respectively. What is now, approximately, the probability that the walk ever hits 10,000? The recursion is now much simpler

\[ p_0 = 1, \]
\[ p_n = p \cdot p_{n-1} + (1 - p) \cdot p_{n-2}, \]

and we can solve it, but again we can avoid all work by applying the theorem to get that \( p_n \) converges to \( \frac{1}{2-p} \).

**Proof.** We can assume, without loss of generality that \( f_N > 0 \) (or else reduce \( N \)).

Define a Markov chain with state space \( S = \{0, 1, \ldots, N-1\} \) by

\[
\begin{bmatrix}
  f_1 & 1-f_1 & 0 & 0 & \cdots \\
  f_2 & 1-f_2 & 0 & 0 & \cdots \\
  \frac{f_1}{1-f_1} & 0 & 1-f_1-f_2 & 0 & \cdots \\
  \frac{f_2}{1-f_1-f_2} & 0 & 0 & 1-f_1-f_2-f_3 & \cdots \\
  \cdots & & & & \\
  \frac{f_N}{1-f_1-\cdots-f_{N-1}} & 0 & 0 & 0 & \cdots \\
\end{bmatrix}
\]

This is called *renewal chain*: it moves to the right (from \( x \) to \( x + 1 \)) on nonnegative integers, except for *renewals*, i.e., jumps to 0. At \( N-1 \), the jump to 0 is certain (note that the matrix entry \( P_{N-1,0} \) is 1, since the sum of \( f_k \)'s is 1).

The chain is irreducible (you can get to \( N-1 \) from anywhere, from \( N-1 \) to 0, and from 0 anywhere) and we will see shortly that is also aperiodic. If \( X_0 = 0 \), and \( R_0 \) is the first return time to 0, then

\[
P(R_0 = k)
\]

clearly equals \( f_1 \) if \( k = 1 \), then for \( k = 2 \) it equals

\[
(1-f_1) \cdot \frac{f_2}{1-f_1} = f_2,
\]

then for \( k = 3 \) it equals

\[
(1-f_1) \cdot \frac{1-f_1-f_2}{1-f_1} \cdot \frac{f_3}{1-f_1-f_2} = f_3,
\]
and so on. We conclude that (recall again that $X_0 = 0$)

$$P(R_0 = k) = f_k \quad \text{for all } k \geq 1.$$ 

In particular, the promised aperiodicity follows, as the chain can return to 0 in $k$ steps if $f_k > 0$. Moreover, the expected return time to 0 is

$$m_{00} = \sum_{k=1}^{N} kf_k = \mu.$$ 

The next observation is that the probability $P_{00}^n$ that the chain is at 0 in $n$ steps is given by the recursion

$$(17.6) \quad P_{00}^n = \sum_{k=1}^{n} P(R_0 = k)P_{00}^{n-k}.$$ 

To see this, observe that you must return sometime by time $n$ in order to end up at 0; either you return for the first time at time $n$, or you return at some previous time $k$, and then you have to be back at 0 in $n - k$ steps.

The above formula (17.6) is true for every Markov chain. In this case, however, we note that for sure the first return time to 0 is at most $N$, so we can always sum to $N$ with the proviso that $P_{00}^{n-k} = 0$ when $k > n$. So from (17.6) we get

$$P_{00}^n = \sum_{k=1}^{N} f_k P_{00}^{n-k}.$$ 

The recursion for $P_{00}^n$ is the same as the recursion for $u_n$. The initial conditions are also the same, and we conclude that $u_n = P_{00}^n$. It follows from the convergence theorem (Theorem 15.3) that

$$\lim_{n \to \infty} u_n = \lim_{n \to \infty} P_{00}^n = \frac{1}{m_{00}} = \frac{1}{\mu},$$

which ends the proof. \qed

Patterns in coin tosses

Assume that you repeatedly toss a coin, with Heads represented by 1 and Tails represented by 0. On any toss 1 occurs with probability $p$. Assume also that you have a pattern of outcomes, say 101101. What is the expected number of tosses needed to obtain this pattern? It should be about $2^7 = 128$ when $p = \frac{1}{2}$, but what is it exactly? One could compare two patterns by this waiting game, saying that the one with smaller expected value wins.

Another way to compare two patterns is the horse race: you and your adversary each choose a pattern, say 1001 and 0100, and the person whose pattern appears first wins.

Here are the natural questions. How do we compute the expectations in the waiting game, and the probabilities in the horse race? Is the pattern that wins in the waiting game more
likely to win in the horse race? There are several ways how to solve this problem (a particularly elegant one uses the so called Optional Stopping Theorem for martingales), but we will be using Markov chains.

The Markov chain $X_n$ we will use has the state space all patterns of length $\ell$. Each time, the chain transitions into the pattern obtained by appending 1 (with probability $p$) or 0 (with probability $1 - p$) at the right end of the current pattern, and deleting the symbol at the left end of the current pattern. That is, the chain simply keeps track of the last $\ell$ symbols in the sequence of tosses.

There is a slight problem before we have $\ell$ tosses. For now, just assume that the chain starts with some particular sequence of $\ell$ tosses, chosen in some way.

The one thing that we can figure out immediately for this chain is its invariant distribution. At any time $n \geq 2\ell$, and any pattern $A$ with $k$ 1’s and $\ell - k$ 0’s,

$$P(X_n = A) = p^k(1 - p)^{\ell - k}$$

as the chain is generated by independent coin tosses! Therefore, the invariant distribution of $X_n$ assigns $A$ the probability

$$\pi_A = p^k(1 - p)^{\ell - k}.$$ 

Now if you have two patterns $B$ and $A$, denote by $N_{B \to A}$ the expected number of additional tosses you need to get $A$ provided that the first tosses ended in $B$. Here, if $A$ is a subpattern of $B$, this does not count, you have to actually make $A$ in the additional tosses, although you can use a part of $B$. For example, if $B = 111001$ and $A = 110$, and next tosses are 10 then $N_{B \to A} = 2$, and if the next tosses are 001110 then $N_{B \to A} = 6$.

Also denote

$$E(B \to A) = E(N_{B \to A}).$$

Our initial example can therefore be formulated as follows: compute

$$E(\emptyset \to 1011101).$$

The convergence theorem for Markov chains guarantees that, for every $A$

$$E(A \to A) = \frac{1}{\pi_A}.$$ 

The hard part of our problem is over. We now show how to analyze the waiting game on the example.

We know that

$$E(1011101 \to 1011101) = \frac{1}{\pi_{1011101}}.$$ 

However, starting with 1011101, we can only use the overlap 101 to help us get back to 1011101, so that

$$E(1011101 \to 1011101) = E(101 \to 1011101).$$
Now to get from ∅ to 1011101 we have to get first to 101, and then from there to 1011101, so that
\[ E(∅ \rightarrow 1011101) = E(∅ \rightarrow 101) + E(101 \rightarrow 1011101). \]

Now we have reduced the problem to 101 and we iterate our method:
\[
E(∅ \rightarrow 101) = E(∅ \rightarrow 1) + E(1 \rightarrow 101) \\
= E(∅ \rightarrow 1) + E(101 \rightarrow 101) \\
= E(1 \rightarrow 1) + E(101 \rightarrow 101) \\
= \frac{1}{π_1} + \frac{1}{π_{101}}.
\]

The final result is
\[
E(∅ \rightarrow 1011101) = \frac{1}{π_{1011101}} + \frac{1}{π_{101}} + \frac{1}{π_1} \\
= \frac{1}{p^5(1-p)^2} + \frac{1}{p^2(1-p)} + \frac{1}{p},
\]
which is equal to \(2^7 + 2^3 + 2 = 138\) when \(p = \frac{1}{2}\).

In general, the expected time \(E(∅ \rightarrow A)\) can be computed by adding to \(1/π_A\) all the overlaps between \(A\) and its shifts, that is, all the patterns by which \(A\) begins and ends. In the example, the overlaps are 101 and 1. The more overlaps \(A\) has, the larger \(E(∅ \rightarrow A)\) is. Accordingly, for \(p = \frac{1}{2}\), of all patterns of length \(ℓ\), the largest expectation is \(2^ℓ + 2^{ℓ-1} + \cdots + 2 = 2^{ℓ+1} - 2\) (for constant patterns 11...1 and 00...0) and the smallest \(2^ℓ\) when there is no overlap at all (for example, for 100...0).

Now that we know how to compute expectations in the waiting game, we will look at the horse race. Fix two patterns \(A\) and \(B\) and let \(p_A = P(A\ \text{wins})\) and \(p_B = P(B\ \text{wins})\). The trick is to consider the time \(N\), the first time one of the two appears. Then we can write
\[ N_{∅→A} = N + I_{B\ \text{appears before } A}N'_{B→A}, \]
where \(N'_{B→A}\) is the additional number of tosses you need to get to \(A\) after you reach \(B\) for the first time. In words, to get to \(A\) you either stop at \(N\) or go further starting from \(B\), but the second case only occurs when \(B\) occurs before \(A\). It is clear that \(N'_{B→A}\) has the same distribution as \(N_{B→A}\) and is independent of the the event that \(B\) appears before \(A\). (At the time \(B\) appears for the first time, what matters for \(N'_{B→A}\) is that you are at \(B\) and not on whether you have seen \(A\) earlier.) Taking expectations,
\[
E(∅ \rightarrow A) = E(N) + p_B \cdot E(B \rightarrow A), \\
E(∅ \rightarrow B) = E(N) + p_A \cdot E(A \rightarrow B),
\]
\[ p_A + p_B = 1. \]

We already know how to compute \(E(∅ \rightarrow A)\), \(E(∅ \rightarrow B)\), \(E(B \rightarrow A)\), and \(E(A \rightarrow B)\) so this is a system of three equations with three unknowns: \(p_A, p_B\) and \(N\).

**Example 17.3.** Let’s return to the patterns \(A = 1001\) and \(B = 0100\), and \(p = \frac{1}{2}\), and compute the winning probabilities in the horse race.
We compute $E(\emptyset \rightarrow A) = 16 + 2 = 18$ and $E(\emptyset \rightarrow B) = 16 + 2 = 18$. Next we compute $E(B \rightarrow A) = E(0100 \rightarrow 1001)$. First we note $E(0100 \rightarrow 1001) = E(\emptyset \rightarrow 1001)$, so that $E(0100 \rightarrow 1001) = E(\emptyset \rightarrow 1001) - E(\emptyset \rightarrow 100) = 18 - 8 = 10$. Similarly, $E(A \rightarrow B) = 18 - 4 = 14$, and then the above three equations for three unknowns give $p_A = \frac{5}{12}$, $p_B = \frac{7}{12}$, $E(N) = \frac{73}{6}$.

We conclude with two examples, each somewhat paradoxical and thus illuminating about probability.

**Example 17.4.** Consider sequences $A = 1010$ and $B = 0100$. It is straightforward to verify that $E(\emptyset \rightarrow A) = 20$, $E(\emptyset \rightarrow B) = 18$, while $p_A = \frac{9}{14}$. So $A$ loses in the waiting game but wins in the horse race! What is going on? Simply, when $A$ loses in the horse race, it loses by a lot, thereby tipping the waiting game towards $B$.

**Example 17.5.** This example is concerned by the horse race only. Consider the relation $\geq$ given by $A \geq B$ if $P(A$ beats $B) \geq 0.5$. Naively, one would expect that this relation is transitive, but this is not true! The simplest example are triples $011 \geq 100 \geq 001 \geq 011$, with probabilities $\frac{1}{2}$, $\frac{3}{4}$ and $\frac{2}{3}$.

**Problems**

1. Start at 0 and perform the following random walk on the integers. At each step, flip 3 fair coins and make a jump forward equal to the number of Heads (you stay where you are if you flip no Heads). Let $p_n$ be the probability that you ever hit $n$. Compute $\lim_{n \to \infty} p_n$. (It is not $\frac{2}{3}$!)  

2. Say you have three patterns $A = 0110$, $B = 1010$, $C = 0010$. Compute the probability that $A$ appears first of the three in the sequence of fair coin tosses.

**Solutions to problems**

1. The size $S$ of the step has p. m. f. given by $P(X = 0) = \frac{1}{8}$, $P(X = 1) = \frac{3}{8}$, $P(X = 2) = \frac{3}{8}$, $P(X = 3) = \frac{1}{8}$. Thus  

$$p_n = \frac{1}{8}p_n + \frac{3}{8}p_{n-1} + \frac{3}{8}p_{n-2} + \frac{1}{8}p_{n-3},$$

and so  

$$p_n = \frac{8}{7} \left( \frac{3}{8}p_{n-1} + \frac{3}{8}p_{n-2} + \frac{1}{8}p_{n-3} \right).$$
It follows that $p_n$ converges to the reciprocal of
\[ \frac{8}{7} \left( \frac{3}{8} \cdot 1 + \frac{3}{8} \cdot 2 + \frac{1}{8} \cdot 3 \right), \]
that is, to $E(S|S > 0)^{-1}$. The answer is
\[ \lim_{n \to \infty} p_n = \frac{7}{12}. \]

2. If $N$ is the first time one of the three appears, we have
\[
E(\emptyset \to A) = EN + p_B E(B \to A) + p_C E(C \to A) \\
E(\emptyset \to B) = EN + p_A E(A \to B) + p_C E(C \to B) \\
E(\emptyset \to C) = EN + p_A E(A \to C) + p_B E(B \to C) \\
p_A + p_B + p_C = 1
\]
and
\[
E(\emptyset \to A) = 18 \\
E(\emptyset \to B) = 20 \\
E(\emptyset \to C) = 18 \\
E(B \to A) = 16 \\
E(C \to A) = 16 \\
E(A \to B) = 16 \\
E(C \to B) = 16 \\
E(A \to C) = 16 \\
E(B \to C) = 16
\]
The solution is $EN = 8$, and $p_A = \frac{3}{8}$, $p_B = \frac{1}{4}$, and $p_C = \frac{3}{8}$. The answer is $\frac{3}{8}$. 